Band-to-Band Tunnel Transistor Design and Modeling for Low Power Applications

Kanghoon Jeon

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2012-86
http://www.eecs.berkeley.edu/Pubs/TechRpts/2012/EECS-2012-86.html

May 10, 2012

Band-to-Band Tunnel Transistor Design and Modeling for Low Power Applications

By

Kanghoon Jeon

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering - Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Chenming Hu, Chair
Professor Tsu-Jae King Liu
Professor Daryl C. Chrzan

Spring 2012

Band-to-Band Tunnel Transistor Design and Modeling for Low Power Applications

by

Kangoon Jeon

Abstract

Band-to-Band Tunnel Transistor Design and Modeling for Low Power Applications

by

Kanghoon Jeon

Doctor of Philosophy in Engineering – Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Chenming Hu, Chair

As the physical dimensions of the MOSFET have been scaling, the supply voltage has not scaled accordingly and thus the power density has been continuously increasing. This is mainly due to the fact that transistor operation requires carriers to go over the source side potential barrier which limits the subthreshold swing of a MOSFET to 60mV/dec at room temperature and thus inhibits the scaling of the threshold voltage. Tunneling devices utilizing the band-to-band tunneling mechanism have been known to overcome this fundamental limit.

In this thesis, the tunneling field-effect-transistor (TFET) is explored to replace conventional MOSFETs for low power applications. The band-to-band tunneling mechanism is looked into in order to develop a more accurate tunneling model that considers the change in effective mass during the transition between the conduction and valence band. Device simulator parameters are modified with this model and are used in designing the TFET. The silicon P-I-N structure TFET is studied through simulation and various experimental splits as a baseline for the TFET development. High tunneling currents are measured from a short channel device with a flash and spike anneal combination and a novel silicided source TFET using silicide induced dopant segregation is shown to achieve sub-60mV/dec subthreshold swing. Measurement and analysis methods of the transistor current and subthreshold swing to verify the TFET are discussed. Lower band gap Ge devices and Strained Si/Ge hetero-structure devices utilizing a lower effective bandgap are also explored to improve the performance of the TFET.

Dedicated to my family

Table of Contents

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor Professor Chenming Hu for all his guidance, support and encouragement. I thank him for his unwavering energy, enthusiasm and interest in my work. It has led me to become a better engineer and researcher. I would also like to thank Professor Tsu-Jae King Liu for serving as the chair in my qualifying exam committee and for all her advice and valuable insight during our group meetings. I would like to thank Professor Daryl C. Chrzan for serving in both my qualifying exam committee and dissertation committee.

I would like to thank all the device group members for all their help and the wonderful experience. In particular I would like to thank Jemin Park and Changhwan Shin for being my closest friends. We started graduate school together and I thank them for their friendship and support throughout our years in Berkeley. I have been fortunate to have worked with great group-mates, Dr. Anupama Bowonder and Dr. Pratik Patel. I thank them for their support, help and friendship.

I would like to thank Dr. Wei-Yip Loh, Dr. Jungwoo Oh, Dr. Chang-Yong Kang, Dr. Casey Smith, Dr. Chanro Park, Dr. Byoung Gi Min, Dr. Prachant Mahji and Dr. Hsing-Huang Tseng for giving me the opportunity to fabricate TFETs at SEMATECH and for all their expertise, advice and help throughout the length of the project. I would like to thank Dr. Chorng-Ping Chang for allowing me the opportunity to experiment at Applied Materials and for all his advice.

I would like to thank the Korea Foundation for Advanced Studies (KFAS) for the financial support during my graduate studies and DARPA STEEP for funding and supporting this research.

Finally, I would like to thank my family for all their love, support and sacrifice. They have supported and encouraged me all throughout my life. I would not be here were it not for my parents and my brother. Lastly but most certainly not least, I would like to thank my lovely wife, Heejung Kim for her devoted love, support and companionship.

Chapter 1

Introduction

1.1 Power Consumption Becoming a Major Concern

As the transistor has been scaling, the power consumption in modern microelectronic circuits have been continuously increasing. As can be seen in Figure 1.1 (a), the active as well as the passive power density is showing an continuous increase with scaling of the gate length. This is becoming a major concern in data centers due to cost increases in powering and cooling the servers. Mobile devices suffer from shorter battery life due to increased power consumption. In order to reduce the power density, the supply voltage has to scale in accordance to physical scaling dimensions which according to classical scaling rules keeps the power density constant.

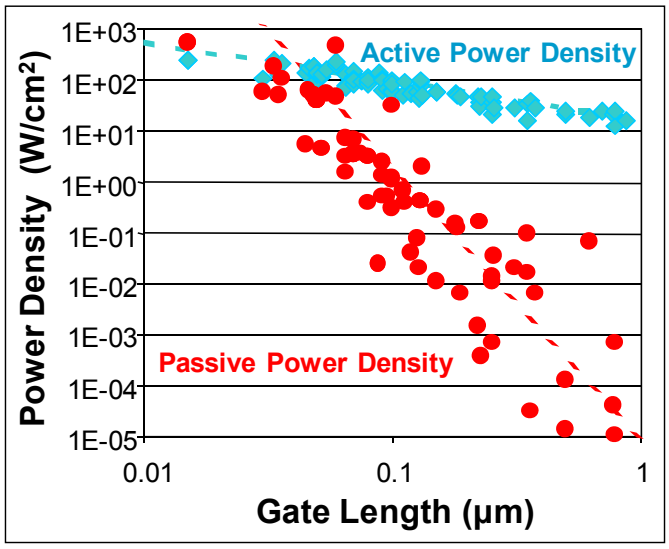


Figure 1.1 Active and passive power density plotted against gate length showing a continuous increase in power density [1.1]

But as can be seen in the following figure, the supply voltage scaling has deviated from the classical scaling since 0.13µm technology generation and shows to be saturating at around 1V. This slowing down of supply voltage scaling has caused the continuous increase in power density.

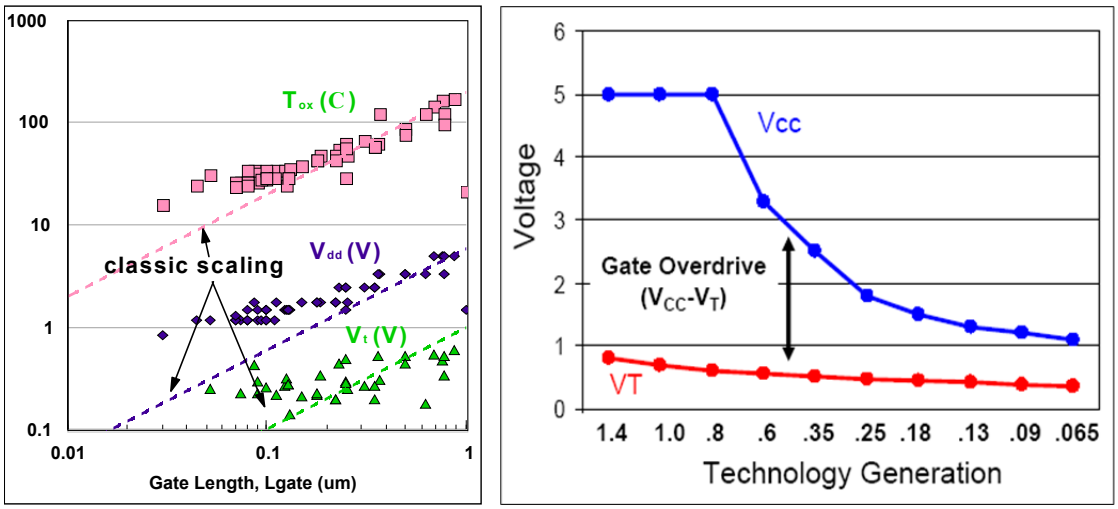


Figure 1.2 **(a)** Scaling of gate T_{OX}, V_{dd} and V_t versus gate length shows deviation from classic scaling [1.1] **(b)** V_{CC} and V_T scaling trend versus technology generation shows a saturation in the voltage scaling [1.2]

The MOSFET operates by having carriers travel over the source-side potential barrier. And since this is a thermal process, it is limited by *kT/q* which corresponds to a subthreshold swing of 60mV/dec. This limit in subthreshold swing has hindered the scaling of the threshold voltage as can be seen in Figure 1.2 and consequently the supply voltage. So in reducing the supply voltage of a MOSFET, a window must be chosen where one would sacrifice the I_{ON} and the other sacrificing I_{OFF} as shown in the following figure.

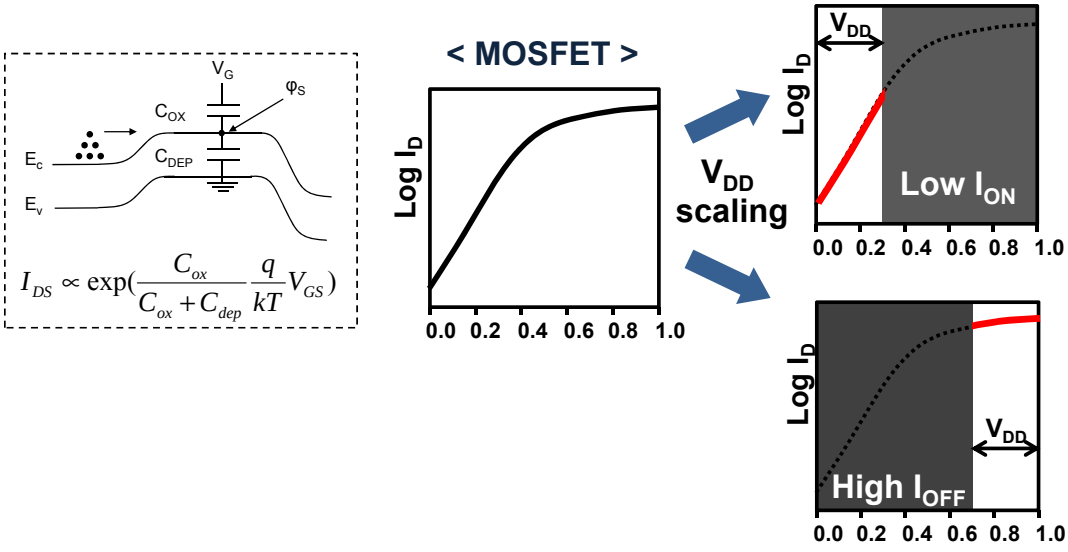


Figure 1.3 Operation of a MOSFET and supply voltage scaling scenarios

1.2 Possible Solution

To achieve good ON/OFF performance at a lower supply voltage, we would have to find a way to scale down the subthreshold swing. That is, overcome the 60mV/dec limit in subthreshold swing of MOSFETs. In order to do this, a new current mechanism that does not involve carriers traveling over a potential barrier needs to be used. The band-to-band tunneling mechanism where electrons in the valence band travel through the barrier to the conduction band has been reported to be not subjected to this limit [1.3]. Tunneling Field-Effect Transistors (TFETs) have been explored by many researchers but have not been able to achieve a comparable performance to the conventional MOSFET [1.3~1.5]. More careful exploration into the structure and fabrication process design of the TFET is required for a low power alternative to the MOSFET.

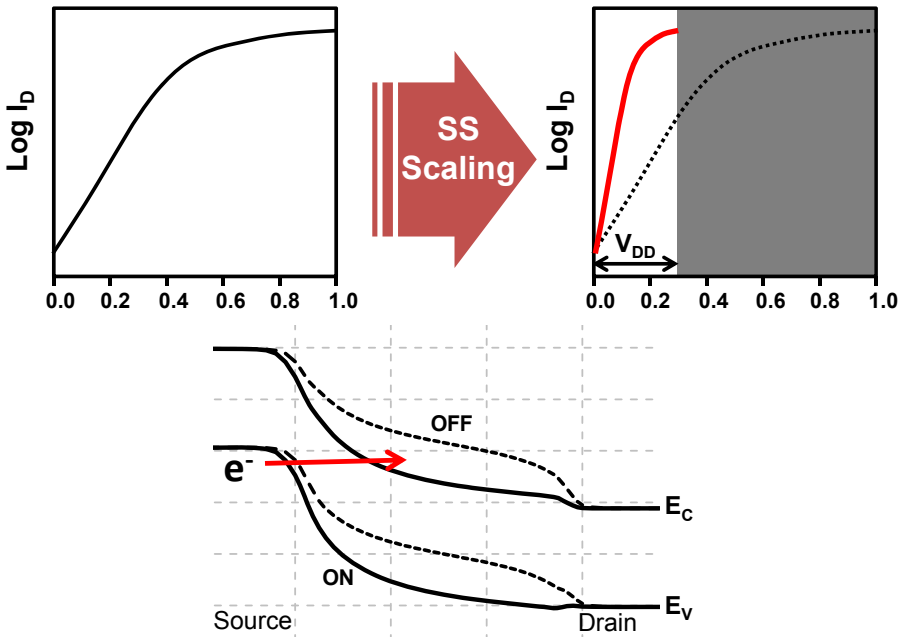


Figure 1.4 Low voltage operation can be achieved by scaling of the subthreshold swing (SS) with band-to-band tunneling

1.3 Thesis Outline

The focus of this research is in exploring the band-to-band tunneling mechanism to be utilized as the main current mechanism of transistors in order to achieve the performance of MOSFETs at a much lower supply voltage with a less than 60mV/dec subthreshold swing. This will allow us to reduce the power consumption of electronic circuits by a considerable amount. In Chapter 2, the band-to-band tunneling model is examined to be applied to device simulation. Chapter 3 explores into the fabrication of TFETs with a baseline process flow development.

Electrical characterization of the TFETs is discussed as well. The subthreshold swing of the TFET is improved using silicidation induced dopant segregation and is dealt with in Chapter 4. Data screening criteria for TFET analysis are proposed [1.6]. In Chapter 5, we discuss structures to higher the drive current of TFETs and show improvements using increased gate-to-source overlap with spike+flash anneal combination [1.7~1.8]. Chapter 6 explores germanium TFET structures in homo- and hetero-structure configurations with silicon to better the performance. It is shown that biaxial strain between silicon and germanium will allow for a very low tunneling bandgap.

1.4 References

[1.1] B. Meyerson, IBM, *Semiconductor Conference*, January 2004

[1.2] Paul Packan, "Device and Circuit Interactions", *International Electron Devices Meeting (IEDM) Short Course*, 2007

[1.3] W. Y. Choi, B.-K. Park, J. D. Lee, and T.-J. King Liu, "Tunneling Field-Effect Transistors (TFETs) With Subthreshold Swing (SS) Less Than 60 mV/dec", *IEEE Electron Device Letters*, vol. 28, no. 8, pp. 743-745, Aug. 2007

[1.4] W. M. Reddick and G. A. Amaratunga, "Silicon surface tunnel transistor," *Applied Physics Letters*, vol. 67, no. 4, pp. 494-496, July 1995

[1.5] C. Aydin, A. Zaslavsky, S. Luryi, S. Cristoloveanu, D. Mariolle, D. Fraboulet, S. Deleonibus, "Lateral interband tunneling transistor in silicon-on-insulator," *Applied Physics Letters*, vol. 84, no. 10, pp. 1780-1782, March 2004

[1.6] K. Jeon, W.-Y. Loh, P. Patel, C.Y. Kang. J. Oh, A. Bowonder, C. Park, C.S. Park, C. Smith, P. Majhi, H.-H. Tseng, R. Jammy, T.-J.K. Liu, and Chenming Hu, "Si tunnel transistors with a novel silicided source and 46mV/dec swing," *VLSI Technology, 2010 Symposium on*, pp. 121-122, June 2010

[1.7] W.-Y. Loh, K. Jeon, C. Y. Kang, J. Oh, P. Patel, C. Smith, J. Barnett, C. Park, T.-J. King Liu, H.-H. Tseng, P. Majhi, R. Jammy and C. Hu, "Sub-60nm Si Tunnel Field Effect Transistors with I_{ON} > 100μA/μm", *Proceedings of ESSDERC*, pp. 162-165, Sept. 2010

[1.8] W.-Y. Loh, K. Jeon, C. Y. Kang, J. Oh, T.-J. King Liu, H.-H. Tseng, W. Xiong, P. Majhi, R. Jammy, and C. Hu, "Highly scaled (L_g ~ 56nm) gate-last Si tunnel field-effect transistors with I_{ON} > 100μA/μm", *Solid-State Electronics*, vol. 65–66, pp. 22-27, Nov.–Dec. 2011

Chapter 2

Modeling of Band-to-Band Tunneling

2.1 Introduction

The band-to-band tunneling phenomenon which was first observed over 50 years ago in narrow bandgap germanium p-n junction diodes by Esaki [2.1], operates by having electrons travel through the forbidden bandgap and thus has been known to be able to overcome the 60mV/dec limit in subthreshold swing of conventional MOSFETs. This interband tunneling mechanism has been used in P-I-N structure TFETs and shown to have a steeper slope [2.2]. In spite of this steep slope, the on current of the TFET is much lower than conventional MOSFETs [2.3~2.4]. In order to achieve an increase in the on current a careful study of the tunneling mechanism is required.

In this chapter, the band-to-band tunneling mechanism is modeled and compared to that of the device simulator MEDICI. We start off with the form of Kane's tunneling model [2.5] where only one effective mass of the carrier is considered. It is expanded to consider the transition of the effective mass from the valence to the conduction band, inside the forbidden gap. Tunneling rates using this model is calculated for silicon and germanium then put into MEDICI for more exact silicon and germanium tunneling simulations. The dependence of the tunneling rate on effective masses is also studied.

2.2 Model Development

The exact form of the wave function throughout the band-to-band transition is required in order to calculate the tunneling probability accurately. Since the wave travels through the forbidden gap, the form of the wave vector inside the gap is required and inside the gap, the wave function undergoes a decay and is purely imaginary. This imaginary wave vector can be found if we consider the law of conservation of energy. If we assume there is no loss of energy

during the transition, the energy of the electron ending up in the conduction band should be equal to the sum of its energy at the start of tunneling in the valence band and the gained energy from the acceleration due to the electric field [2.6]. This is shown in the following equation,

$$\left(E_v - \frac{\hbar^2 k^2}{2m_v^*}\right) + qFx = E_c + \frac{\hbar^2 k^2}{2m_c^*} \tag{2.1}$$

where F denotes the electric field and m_c^* and m_v^* are valence and conduction band effective masses. Here we have assumed the momentum is conserved as well which means that there is no scattering involved. So this equation would not be exactly accurate for tunneling in indirect band-gap materials but for simplicity we do not consider the phonon interactions here. The indirect tunneling probability can be considered into the equation by adding a factor to the end of the equation for the phonon occupation number and probability of the scattering event. Another assumption that has been made in this equation is that the electric field is considered to be constant. Although in the actual case the field would not be constant throughout the transition, using the value of the average field agrees well with the path integral values [2.7].

Using (2.1) we can obtain the expression for the wave vector.

$$k = \frac{i}{\hbar}\sqrt{2m_r\left(E_g - qFx\right)} \tag{2.2}$$

where $E_g = E_c - E_v$ and m_r is the reduce effective mass and is expressed as

$$\frac{1}{m_r} = \frac{1}{m_c^*} + \frac{1}{m_v^*} \tag{2.3}$$

Equation (2.2) works well within the energy conservation framework considered when deriving (2.1), but if we calculate the values of k throughout the transition from the valence band to conduction band, we can see that the equation has some flaws. Since the expression is purely imaginary it should reduce to zero at the band edges, that is at the beginning and end of the tunneling process in order to allow a smooth change in the wave vector. The imaginary wave vector does go to zero at the end of the tunneling process if we assume an average electric field which would be the band bending (E_g) divided by the distance. But at the beginning of tunneling ($x = 0$), the imaginary wave vector holds a large value. This cannot be true since the imaginary parts of the wave vector are zero in the valence band and this calls for a sudden jump in the imaginary wave vector. In order to get rid of this sudden jump in the imaginary wave vector, we consider a parabolic shape barrier for the band-to-band tunneling process.

$$k = i\frac{\sqrt{2m_r}}{\hbar}\sqrt{(qFx)\left(1 - \frac{qFx}{E_g}\right)} \tag{2.4}$$

The barrier which was triangular ($E_g - qFx$) in the previous equation is now parabolic $(qFx)\left(1 - qFx/E_g\right)$ and reduces to zero at the beginning and end. The wave vector now shows a smooth transition at the band edges.

Equations (2.2) and (2.4) both consider the different effective masses in the conduction and valence band but during the transition, they lump the two masses into one reduced effective mass m_r and consider it constant throughout the whole tunneling process. This may not be an issue when the conduction and valence band effective masses are similar in value but when they differ substantially in value such as the case for InAs where $m_c^* = 0.023m_0$ and $m_v^* = 0.41m_0$ for heavy holes, the reduced effective mass will not model the transition accurately. Here m_0 denotes the free electron mass. The tunneling electrons will behave more like electrons in the valence band when it is closer the valence band and like electrons in the conduction band when closer to the conduction band. So in order to model this transition of the effective masses, we use a linearly interpolated expression of the energies in each band as it transitions from the valence band to the conduction band [2.8].

$$\frac{\hbar^2 k_x^2}{2m_v^*} - \left(\frac{\hbar^2 k_x^2}{2m_v^*} - \frac{\hbar^2 k_x^2}{2m_c^*}\right)\frac{E}{E_g} + E_\perp = -E\left(1 - \frac{E}{E_g}\right) \tag{2.5}$$

Here, the direction of tunneling is chosen to be in the x direction, so k_x is used and the transverse energy is also added to consider the fact that not all the energy will contribute to the kinetic energy involved in the tunneling process and we assume the transverse energy is conserved. E denotes the energy gained from the electric field caused by the band bending which is qFx. Using this equation, the energy reduces to the energy in the valence band at the beginning of tunneling which only leaves m_v^* in the equation and to the energy in the conduction band at the end of tunneling leaving only m_c^*. This gives a more smooth and natural transition in the energies at the band edges. With this expression we can obtain a more accurate form of the wave vector inside the forbidden gap given as follows.

$$k_x = i\frac{\sqrt{2m_v^*}}{\hbar}\sqrt{\frac{E\left(1 - \frac{E}{E_g}\right) + E_\perp}{\left(1 - \left(1 - \frac{m_v^*}{m_c^*}\right)\frac{E}{E_g}\right)}} \tag{2.6}$$

Figure 2.1 compares the imaginary wave vector of InAs using equation (2.4) and (2.6). It can be seen that when the transition between the masses are not considered, the wave vector shows a maximum at the exact center of the bandgap but with the transition considered, the maximum shifts towards the band with the smaller effective mass. This shows that the two effective masses have different contributions to the imaginary wave vector with the larger hole effective mass dominating in this case.

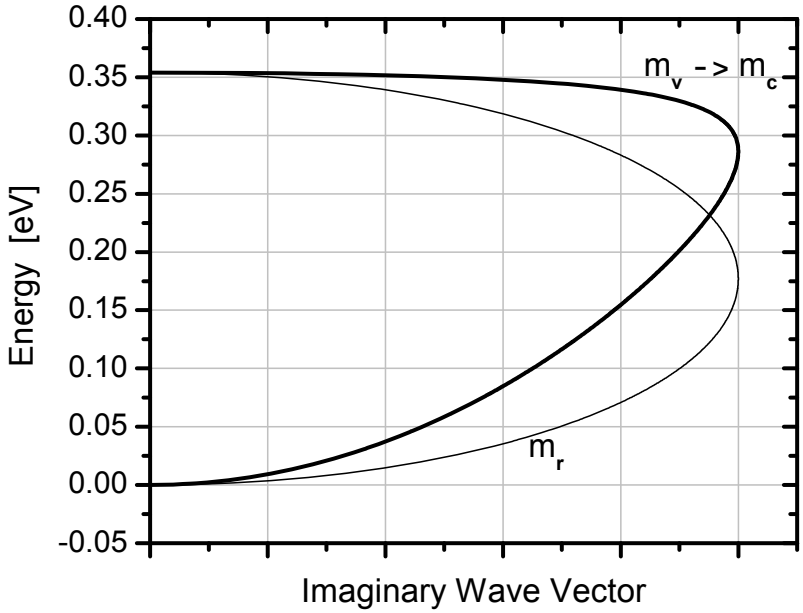


Figure 2.1 Imaginary wave vector (normalized units) vs. energy inside the forbidden bandgap showing the asymmetric transition due to the large difference between the electron and hole effective masses

Now if we use the WKB (Wentzel-Kramers-Brillouin) approximation, we can calculate the tunneling probability. Here we assume a slowly varying potential compared to the wavelength of the electron. The tunneling probability is given as follows from the WKB approximation [2.9],

$$T \cong exp\left(2i\int_{x_1}^{x_2} k_x dx\right) \tag{2.7}$$

where x_1 and x_2 are the classical turning points representing the beginning and ending position of the band-to-band tunneling process. Here we can plug in the wave vector obtained in equation (2.6) to calculate the one dimensional tunneling probability.

$$T \cong exp\left(-2\int_{x_1}^{x_2} \frac{\sqrt{2m_v^*}}{\hbar} \sqrt{\frac{E\left(1-\frac{E}{E_g}\right)+E_\perp}{\left(1-\left(1-\frac{m_v^*}{m_c^*}\right)\frac{E}{E_g}\right)}}\, dx\right) \tag{2.8}$$

The transverse energy term is included to consider the energies in the other two dimensions which in the calculation of the tunneling probability increases the tunnel barrier and the probability quickly reduces to zero at large transverse energies. Due to this transverse energy term, an analytical form of the tunneling probability cannot be obtained and needs to be calculated numerically. Figure 2.2 shows the calculated Zener generation rate for silicon and germanium plotted against the average electric field.

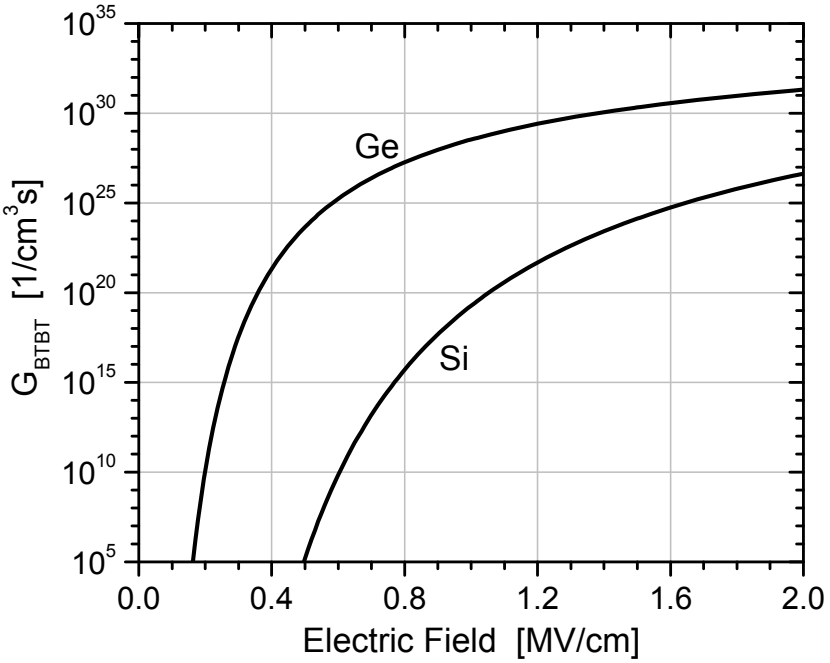


Figure 2.2 Calculated Zener generation rate is plotted against electric field for silicon and germanium. Higher tunneling rate is expected with a lower bandgap material at lower fields but lower bandgap material shows an earlier saturation.

It can be seen that germanium with a lower bandgap, shows a larger Zener generation rate across the entire field range compared to silicon, especially in the low field regime. So in order to achieve high tunneling currents at low voltages, a lower bandgap material would be favorable. But a careful choice in the material is needed with the operation voltage and electric field in mind. For example, InAs with a direct bandgap of 0.36eV has a low density of states and thus shows to have a lower Zener generation rate at high fields compared to germanium with a bandgap of 0.66eV [2.10]. Therefore, depending on the operation voltage, germanium can be a more favorable material compared to InAs for TFETs.

Also in order to compare the performance of each material when used in a TFET, we need to compare the tunneling current for each material with equal tunneling distance. By integrating the tunneling probability and the velocity of the generated electrons in the whole momentum space, we can calculate the tunneling current.

$$J_{btbt} = q \iiint \frac{2}{h^3} |T|^2 v_x dp_x dp_y dp_z \tag{2.9}$$

Using the tunneling probability obtained in equation (2.8), we end up with the following tunneling current density equation.

$$J_{btbt} = \frac{4\pi q m_v^*}{h^3}(qV_a)\int exp\left(-2\int_0^d \frac{\sqrt{2m_v^*}}{\hbar}\sqrt{\frac{(qFx)\left(1-\frac{qFx}{E_g}\right)+E_\perp}{\left(1-\left(1-\frac{m_v^*}{m_c^*}\right)\frac{qFx}{E_g}\right)}}dx\right)dE_\perp \qquad (2.10)$$

An analytic form of the double integral cannot be obtained as well since the transverse energy term is inside the exponent. The tunneling current is calculated numerically using Newton summation (Figure 2.3). The current was calculated for a tunneling width of 10nm.

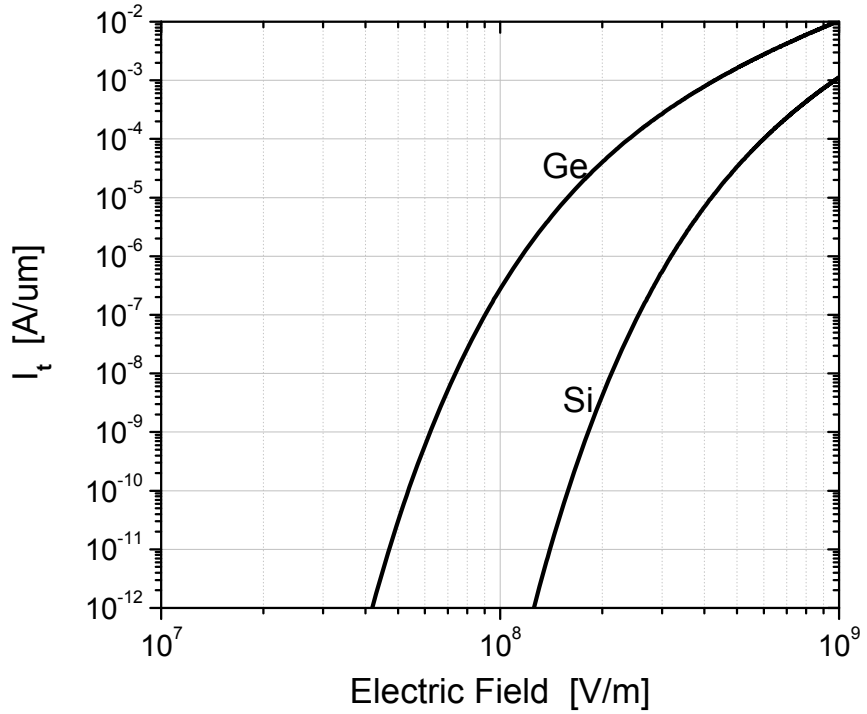


Figure 2.3 Tunneling current vs. electric field for silicon and germanium. The current was calculated for a tunneling width of 10*nm*.

Comparing with Figure 2.2, it can be seen that tunneling current does not increase in germanium as much as the Zener generation rate increases. This is due to the lower density of states in the lower bandgap material. The tunneling probability increases substantially but the resulting tunneling current is not increases as much due to the reduction in the number of carriers that can tunnel through the barrier, since the current is affected by the product of the two factors.

We can see that to achieve 1*mA/μm*, it is practically impossible with silicon since it requires an electrical field of 10^9 *V/m*. This corresponds to a 1V drop in 1nm. But with germanium, we can obtain this amount of current at a more reasonable range of electric field of $2\sim3\times10^8$ *V/m* by increasing the tunneling width to 20*nm*, since the tunneling current increases linearly with the tunneling width.

2.3 Comparison to Simulation

The model used in the device simulator MEDICI has the form of Kane's tunneling model. [2.11]

$$G_{BTBT} = A.BTBT \times \frac{F^{C.BTBT}}{E_g^{1/2}} \times exp\left(-B.BTBT \times \frac{E_g^{3/2}}{F}\right) \tag{2.11}$$

The parameters *A.BTBT*, *B.BTBT* and *C.BTBT* can be defined by the user. Only one set of default values are given which are 3.5×10^{21} $eV^{1/2}/cm\cdot s\cdot V^2$, 22.5×10^{6} $V/cm\cdot eV^{1/2}$ and 2.0 respectively. The Zener generation rate calculated with this model is plotted against the theoretically calculated rates for silicon in Figure 2.4.

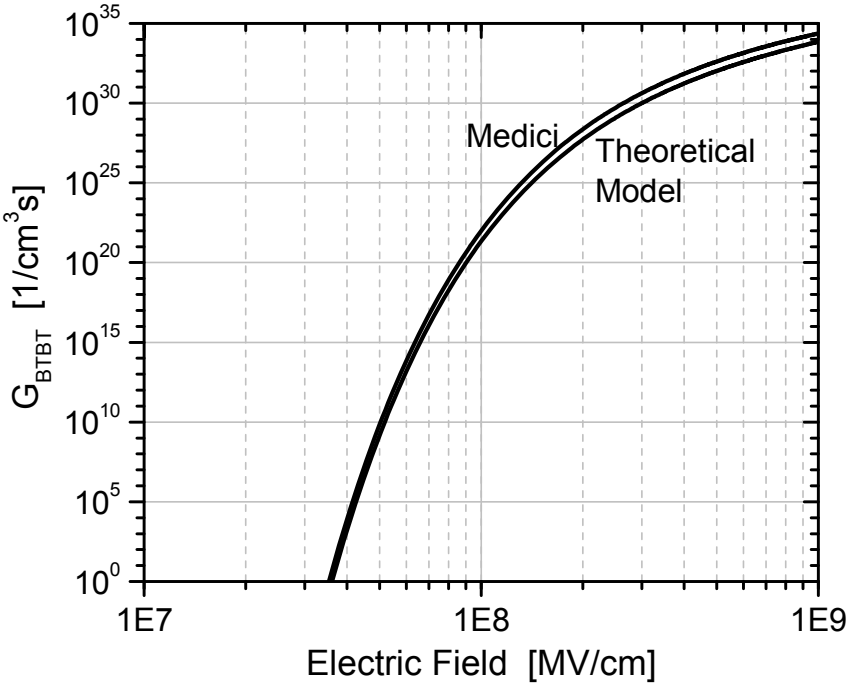


Figure 2.4 Zener generation rate model used by MEDICI and calculated values for silicon. The model agrees well with the calculated results for silicon

The tunneling rate at lower fields agrees fairly well with the theoretically calculated values but at electric fields of higher than 10^8 *V/m*, MEDICI shows a higher tunneling probability. If we consider that the theoretical values were calculated for direct bandgap tunneling, the actual tunneling rate will be even lower than the calculated value.

MEDICI's tunneling model given in (2.11) has a bandgap dependence term to be changed if different materials with different bandgaps are to be simulated. But the simulator only gives one set of parameter values and these values will not be correct since the tunneling characteristics are not only dependent on the bandgap. Figure 2.5 shows the tunneling rate of germanium calculated using the two models.

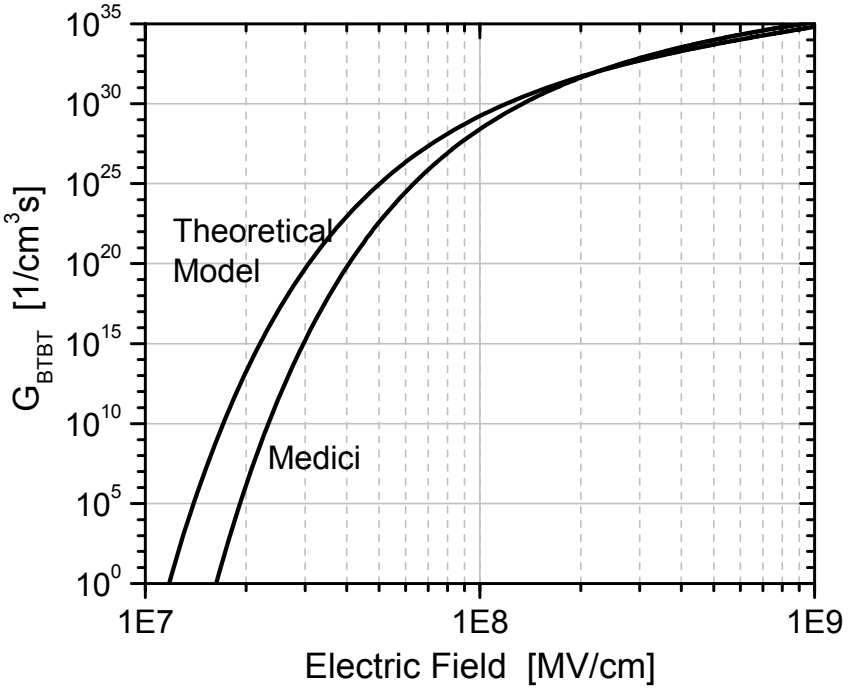


Figure 2.5 Zener generation rate of germanium is plotted for the two models. There is a large difference in the lower field regions.

In the case of germanium a large difference in the rates can be observed in the lower field regions. At higher field values the two curves both saturate to similar tunneling rates but the theoretically calculated value shows orders of magnitude larger tunneling rates at fields below 10^8 *V/m*. The tunneling parameters in MEDICI were changed to fit the calculated rate curve. They were found to be $A.BTBT=9.5\times10^{20} eV^{1/2}/cm\cdot s\cdot V^2$, $B.BTBT=16.7\times10^6$ $V/cm\cdot eV^{1/2}$ and $C.BTBT=2.0$. Using these new parameters, a p-i-n structure TFET structure with a p+ source and n+ drain was simulated. The drain current vs. gate bias for the case using the default parameters and that using the new parameters are plotted in Figure 2.6.

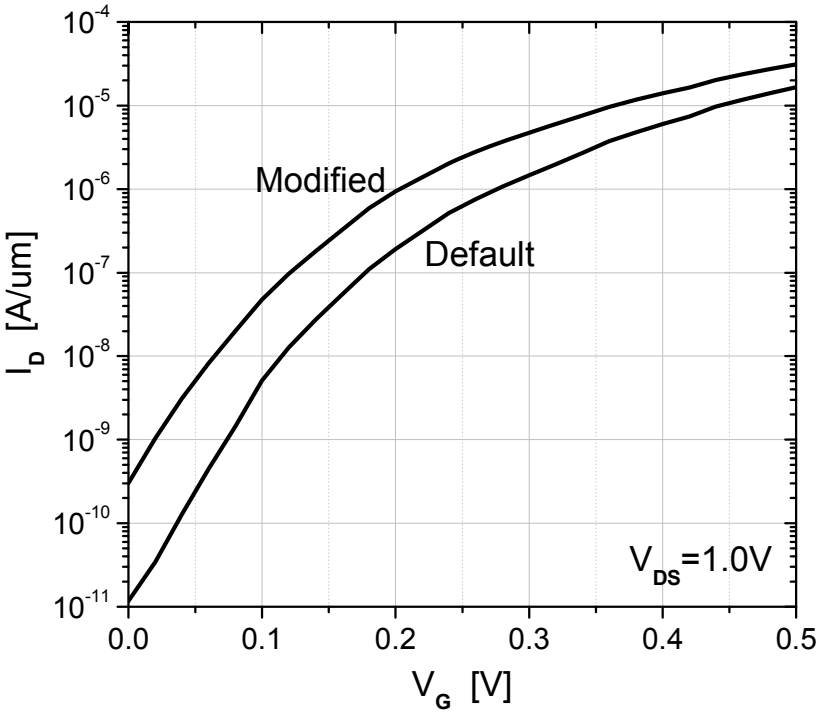


Figure 2.6 I_D vs. V_G for a simple germanium p-i-n TFET structure. MEDICI shows lower tunneling currents in germanium.

As expected, the current obtained using new parameters show a higher on current. But due to the higher tunneling rate at lower fields, an increase in the off state current is also observed. It is shown that only the change of bandgap does not correctly model the tunneling characteristics.

Since most of the tunneling occurs only one direction, in the lateral direction for p-i-n structure TFET's and vertically in pocket type structures, the tunneling electrons will only be subject to one kind of mass. That is, direction dependence in the effective mass also needs to be considered. In order to see how much the tunneling changes according to the directional masses, the tunneling probability from the heavy and light hole mass to longitudinal and transverse electron mass are calculated for each case.

The tunneling from the light hole to the transverse electron mass shows the highest rates since they are the smaller of the two. And accordingly the lowest generation rate is observed in the case of the heavy hole and longitudinal electron effective mass. The two rates differ by more than 10 orders in the lower field regions and approximately 4 orders at an electric field value of 2×10^8 *V/m*. The tunneling current can be increased a substantial amount by designing the device to have the dominant tunneling from the light hole in the direction of the traverse electron effective mass. This can be achieved by using strain to lift the light hole band above the heavy hole band to make the light hole tunneling to occur earlier than heavy holes. And the transverse mass dependency can be achieved by making the tunneling direction to be in the [100] crystal orientation direction for silicon.

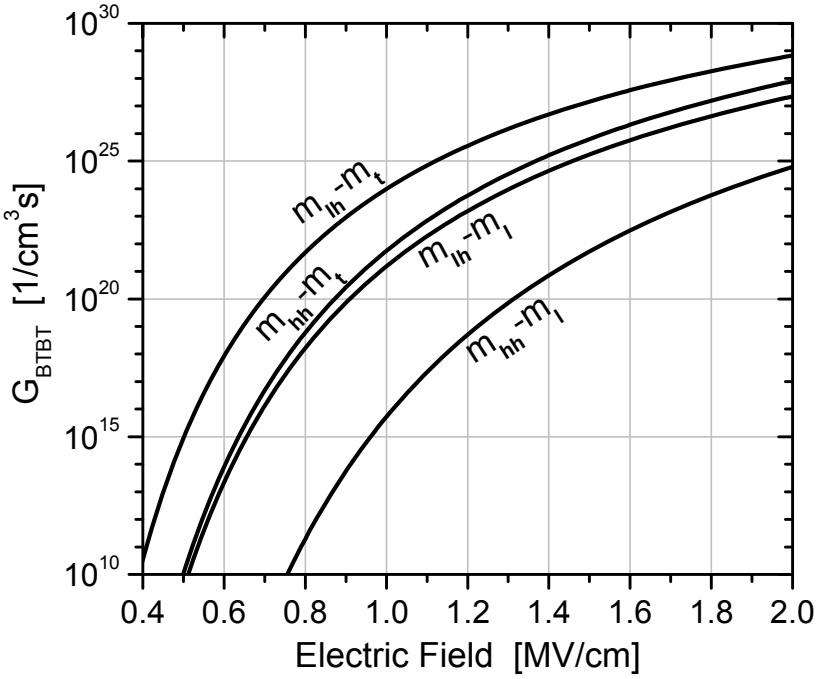


Figure 2.7 Zener generation rate from heavy and light hole mass to longitudinal and traverse effective mass

2.4 Summary

Band-to-band tunneling phenomenon was modeled considering the transition of the effective masses in the two bands. When this transition is not considered, the imaginary wave vector undergoes a parabolic shape change in the forbidden gap with the peak happening at the center of the forbidden bandgap regardless of the differences in effective masses. Whereas when the transition in the effective masses is considered, the peak is shifted towards the band with the heavier effective mass. The imaginary wave vector is calculated by using a linear interpolation of the energies as it transitions from the valence band to conduction band.

The band-to-band tunneling generation rate is calculated using this wave vector and compared to the tunneling parameters in MEDICI. Default silicon tunneling parameters show a good match to the calculated tunneling rate at low field regions but overestimate at high field regions. As for germanium, default parameters are the same as silicon except for the bandgap and thus give a much lower tunneling generation rate at lower fields compared to calculated values. Germanium TFET simulations are performed with modified tunneling parameters and are shown to be able to achieve higher tunneling currents than predicted using default parameters.

Effective mass dependent tunneling generation rate calculations show that by designing the TFET structure so that the tunneling occurs from the light hole band towards the direction where the transverse electron effective mass dominates, an increase in the tunneling current can be achieved.

2.5 References

[2.1] L. Esaki, "New Phenomenon in Narrow Germanium p-n Junctions", *Physical Review*, vol. 109, pp. 603-604, 1958

[2.2] W. Y. Choi, B.-K. Park, J. D. Lee, and T.-J. King Liu, "Tunneling Field-Effect Transistors (TFETs) With Subthreshold Swing (SS) Less Than 60 mV/dec", *IEEE Electron Device Letters*, vol. 28, no. 8, pp. 743-745, Aug. 2007

[2.3] W. M. Reddick and G. A. Amaratunga, "Silicon surface tunnel transistor," *Applied Physics Letters*, vol. 67, no. 4, pp. 494-496, July 1995

[2.4] C. Aydin, A. Zaslavsky, S. Luryi, S. Cristoloveanu, D. Mariolle, D. Fraboulet, S. Deleonibus, "Lateral interband tunneling transistor in silicon-on-insulator," *Applied Physics Letters*, vol. 84, no. 10, pp. 1780-1782, March 2004

[2.5] E. O. Kane, "Theory of Tunneling," *Journal of Applied Physics*, vol. 32, no. 1, pp. 83-91, Jan. 1961

[2.6] S. Wang, Fundamentals of Semiconductor Theory and Device Physics, pp. 481-492, Prentice Hall, New Jersey, 1989

[2.7] P. Patel, "Steep Turn On/Off "Green" Tunnel Transistors," Ph.D. thesis, Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, pp. 16, 2010

[2.8] P. V. Dressendorfer, "Interface and Electron Tunneling Properties of Thin Oxides on Silicon," Ph.D. thesis, Department of Engineering and Applied Sciences, Yale University, pp. 235, 1978

[2.9] D. J. Griffiths, Introduction to Quantum Mechanics, 2nd Edition, pp. 315-339, Prentice Hall, New Jersey, 2005

[2.10] M.V. Fischetti, T.P. O'Regan, S. Narayanan, C. Sachs, J. Seonghoon, J. Kim, Y. Zhang, "Theoretical study of some physical aspects of electronic transport in nMOSFETs at the 10-nm gate-length," *IEEE Transactions on Electron Devices*, vol. 54, no. 9, pp. 2116-2136,
Sept. 2007.

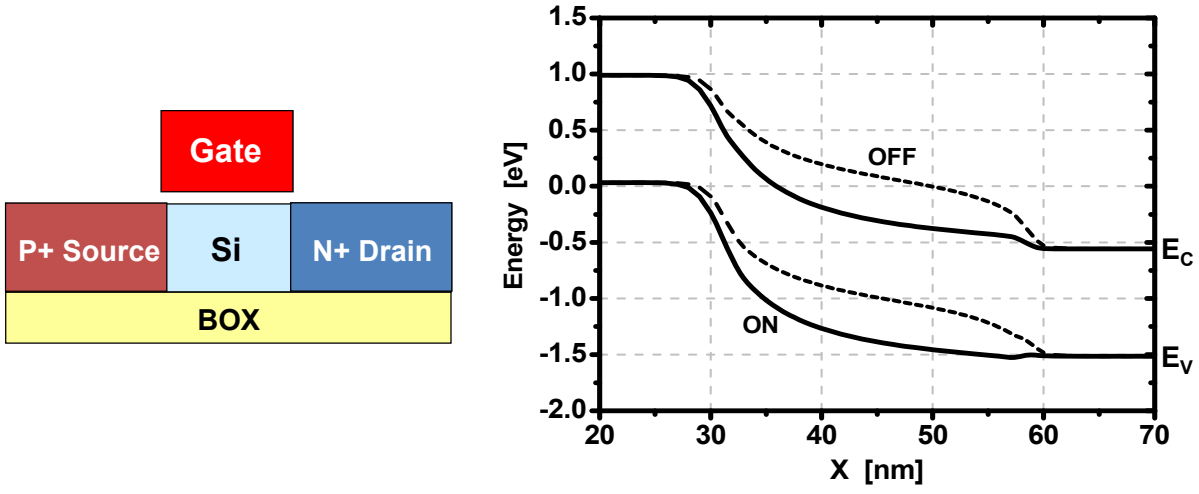
[2.11] Taurus MEDICI User Manual, Version Y-2006.06, Synopsys, Mountain View, CA

Chapter 3

Silicon P-I-N Structure TFET

3.1 Introduction

The simplest TFET is a gated P-I-N diode where the source and drain are highly doped with the gate controlling the band-to-band tunneling between the I-channel region and the P+ or N+ region by way of energy band bending in the I-channel region [3.1~3.5]. Figure 3.1 (a) shows an N-type P-I-N structure TFET. The gate induces an N+ channel to form at the surface of the I-channel region in this case and causes a P+/N+ junction to form at the source to channel interface. Figure 3.1 (b) shows the band diagram of the P-I-N structure TFET close to the surface.



(a) **(b)**

Figure 3.1 **(a)** Illustration of a P-I-N structure TFET (N-type) **(b)** Band diagram of the P-I-N structure TFET close to the surface showing the ON and OFF state of the device.

When the device is in the OFF state, the tunneling path for the electrons from the valence band of the source to the conduction band of the channel is cut off. As a positive gate bias is applied, the gate electric field pulls the band downward which causes the conduction band in the channel to overlap with the valence band of the source. When this overlap occurs the electrons in the valence band of the source are able to tunnel into the conduction band of the channel allowing a sudden increase in the drain current. And since the carriers are not flowing over a potential barrier the subthreshold swing is not subject to the 60mV/dec limit. But in order to achieve sub-60mV/dec, the tunnel barrier needs to be carefully designed. The thickness of the barrier that the electrons have to tunnel through corresponds to the integration range in the tunneling current equation derived in the previous chapter. And since the tunneling current is exponentially dependent on this integral, the output characteristics of the TFET can vary drastically depending on the design of the tunnel barrier. Figure 3.2 shows the simulation results of a 1.0µm channel length P-I-N silicon N-type TFET where the P+ source doping profile is varied using the device simulator MEDICI [3.6]. With the source doping gradient (lateral) changing from 1nm/dec to 20nm/dec, we can see in the I_D-V_G comparison that a large reduction in the tunneling current and degradation in the subthreshold swing. Also, it can be seen that in order to achieve a subthreshold swing of 30mV/dec, a 1nm/dec lateral source doping gradient is required.

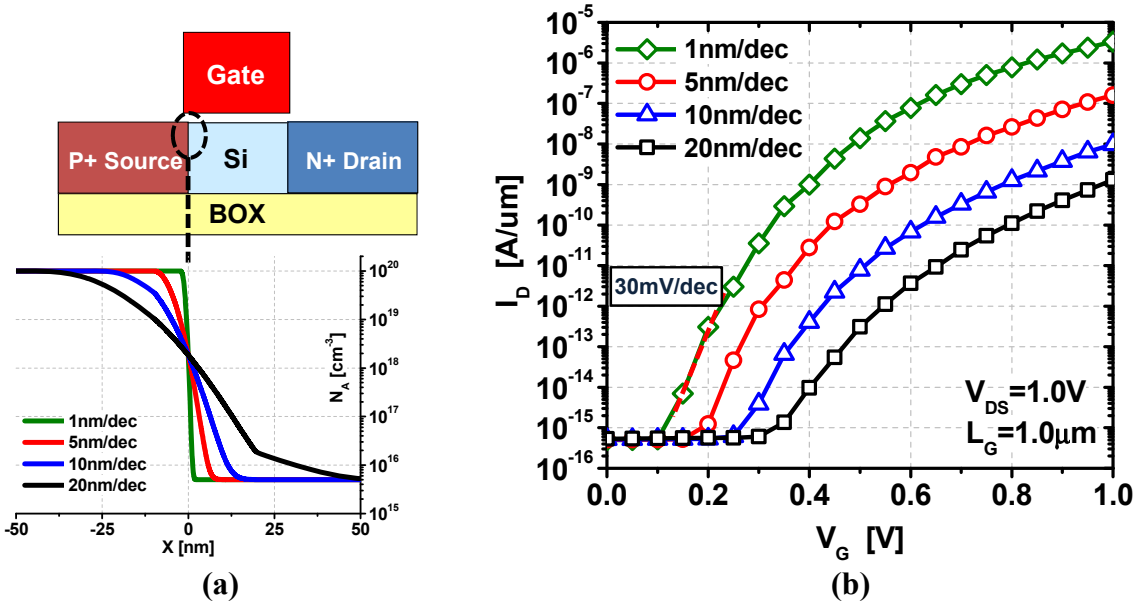


Figure 3.2 **(a)** Simulated P-I-N structure N-type TFET with varying source doping profile **(b)** I_D-V_G curve of the P-I-N TFET showing a 1nm/dec source doping profile is required to achieve 30mV/dec subthreshold swing.

The fabrication of the P-I-N structure TFET is explored in this chapter. First we look into building a process flow for the P-I-N TFET. Process issues observed during the fabrication of the device are also mentioned and solutions to these process issues are given. Experimental results are given and various splits are compared in the analysis.

3.2 Silicon P-I-N Structure TFET Process Flow Setup

Since the P-I-N structure is the simplest TFET, this structure was chosen as the first experimental setup to study the TFET and be used as a baseline for subsequent structures to be fabricated. Due to the asymmetric nature of the P-I-N structure TFET, a self-aligned process similar to the process of a conventional MOSFET is not possible and the source and drain have to be formed separately. The P-I-N TFET process flow is developed using SOI wafers based on Sematech's CMOS process flow using a high-k metal gate process. The process flow is shown in the following figure.

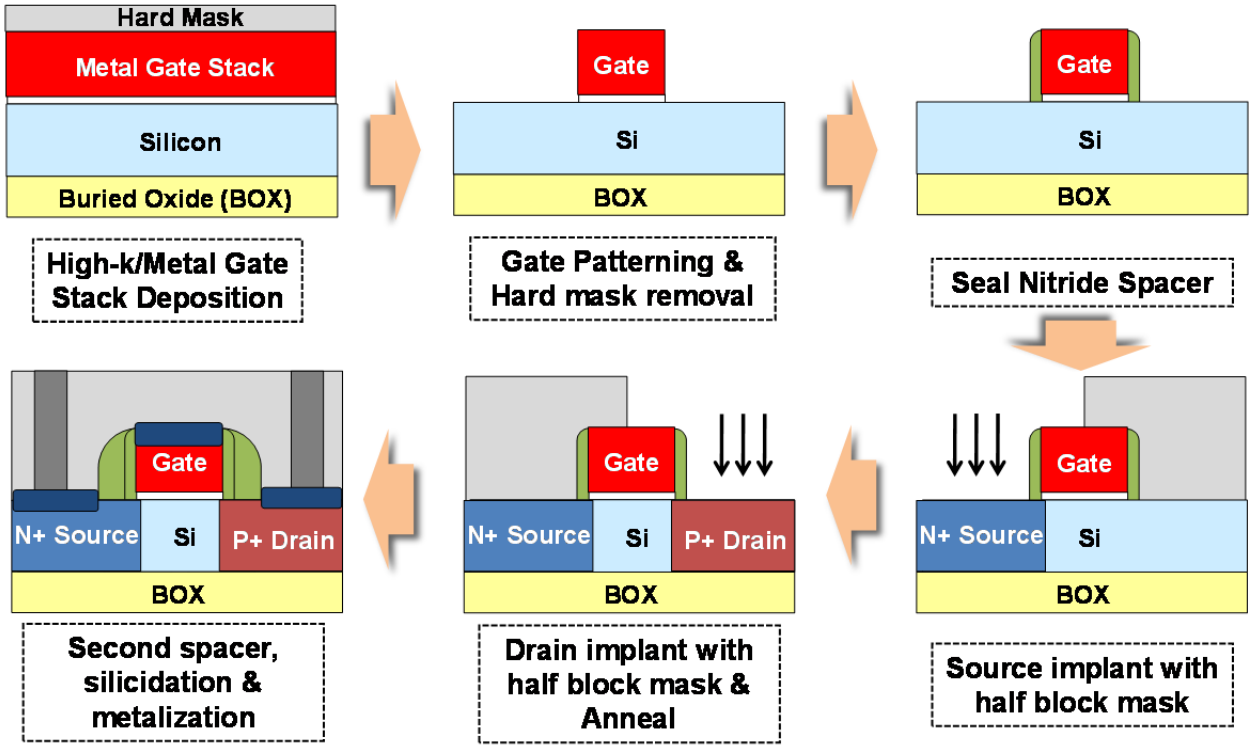


Figure 3.3 P-I-N structure TFET process flow using CMOS mask set

First the SOI wafers were thinned down to 80nm and 40nm using thermal oxidation. Different thicknesses were chosen to see the effect of the SOI thickness to the TFET performance. Mesa isolation was chosen as the method for isolation and active region formation due to its simplicity and less number of steps compared to shallow trench isolation. Then the high-k gate stack is deposited with a mid-gap work function metal and poly-silicon as the gate. After the gate is patterned, a seal nitride spacer is formed to protect the sidewalls of the gate followed by source/drain implant. The source and drain are implanted using a half mask that is compatible with the CMOS mask set and implant conditions were chosen to that the source/drain junction will reach the bottom of the SOI substrate to reduce the junction area and thus limit the amount of junction leakage. This reduces the source/drain junction leakage by more than 2

orders of magnitude. Implant energy and dose splits were included as splits and the effect on the subthreshold swing were examined and will be discussed in the next section. Annealing of the dopants was done using spike anneal at 1050°C for 2 seconds or 5 seconds. Millisecond flash anneal was used as well in some splits. Finally after a second spacer is formed, self-aligned silicide process is done with nickel followed by metallization to M1 layer. Some lots were processed up to silicide without metallization and measured for a quicker turn around.

When mesa isolation is used for active definition on SOI wafers eliminating gate material stringers on the sidewalls of the mesa is an issue. This causes high gate leakage currents due to the increased overlap of the gate to source and drain regions. Also, the exposed regions of the gate dielectric on the sidewalls of the mesa get attacked during the subsequent cleans and implant making the dielectric weaker. And in some cases the poly gate stringer shorts with the source/drain region during silicide process. This phenomenon was observed in the first few lots that were fabricated using mesa isolation. Cross-section SEM images at the edge of active regions showing gate stack material stringers formed on the sidewalls are given in Figure 3.4. High gate currents were measured which caused the analysis of the tunneling current to be difficult. In some cases, we cannot tell whether the drain current is coming from the band-to-band tunneling or from the gate leakage. And when extremely high gate leakages are observed, the drain current showed to be flowing in the opposite direction of the drain bias which shows all the drain current is coming from the gate. Analysis and screening criteria of the tunneling current when gate currents are present are discussed in detail in Chapter 4.

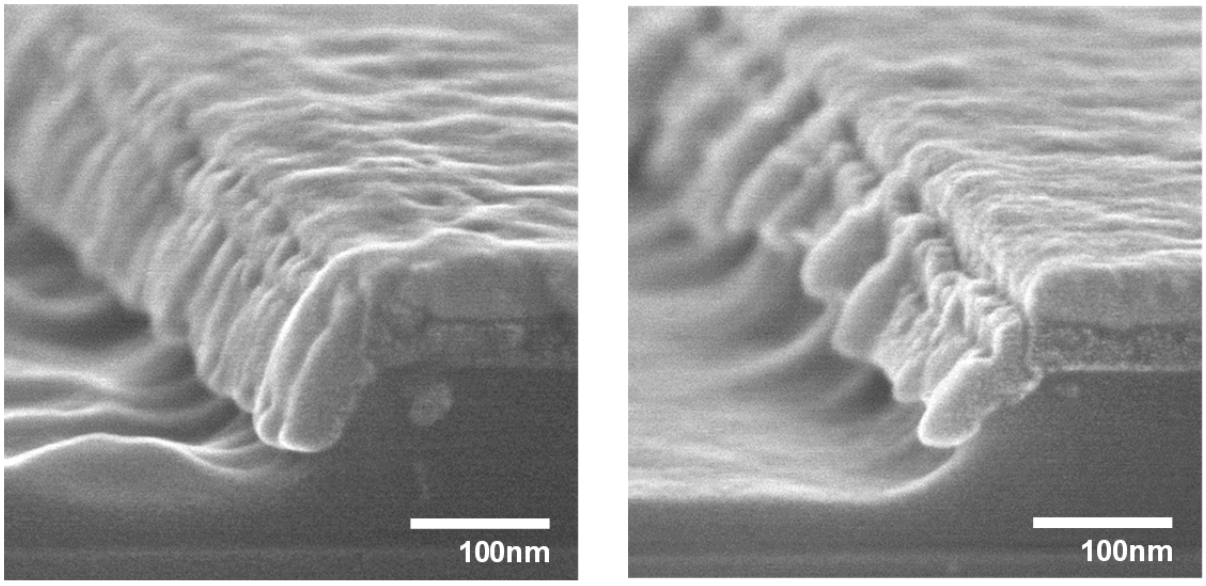


Figure 3.4 Cross-section SEM images showing gate material stringers formed on the sidewalls of the active regions.

In order to prevent from this happening in subsequent lots, the isolation method had to be changed. First shallow trench isolation that is used as a baseline in bulk wafers were considered

but this would require some process module development for it to be applied to SOI wafers and also increase the number of process steps in the isolation to be more than double of that required for the mesa isolation. Deliberately placing sidewall spacers on the active mesa using nitride or oxide was also considered but was not chosen due to the increase in number of steps.

The solution chosen to eliminate this problem was LOCOS (LOCal Oxidation of Silicon). Using the LOCOS process, no process development was required and the increase in number of steps for the isolation was minimal. With the active etch changed to stop on silicon, only a thermal oxidation step was needed to be added to the process flow. A thermal oxide of 1000Å and 2000Å were grown for 40nm and 80nm of substrate thicknesses respectively. Figure 3.5 shows the cross-section SEM images of the active SEM patterns after the LOCOS isolation process. The LOCOS causes the width of the transistor to be reduced by a few 100nm which will have to be considered in small width devices. Also, the gate dielectric thickness increases at the active edges in the width direction due to the bird's beak of the LOCOS. But this actually helps in the operation of the TFET by eliminating early turn-on of the device at the edges from electric field concentration at the corner regions which would cause the subthreshold swing of the TFET to be less sharp.

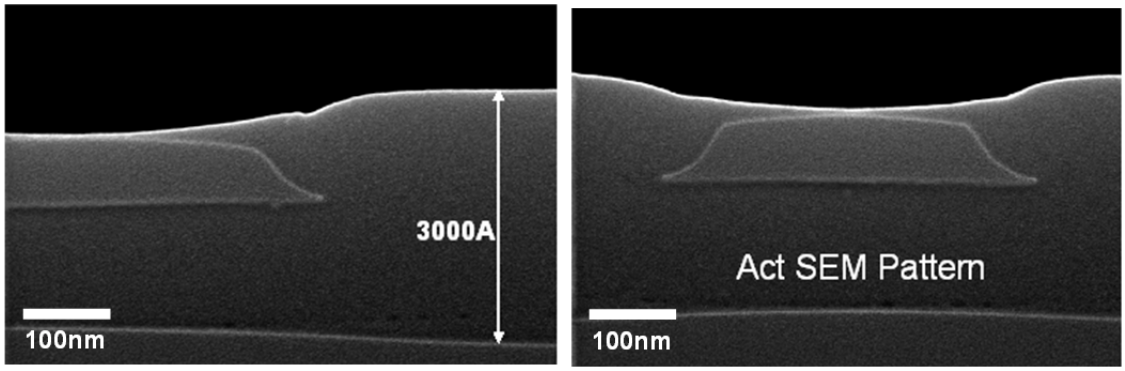


Figure 3.5 Cross-section SEM images of active region after LOCOS isolation process

In addition to the LOCOS isolation, the gate stack etch process was changed to reduce the gate leakage even further. Kang showed that when the high-k gate dielectric is etched after a spacer is formed leaving a protruded high-k region under the gate metal and polysilicon, the gate leakage current can be reduced by reducing the edge leakage [3.7]. As shown in Figure 3.6, the gate leakage increases when recessed and decreases when protruding from the gate edge. The gate leakage shows to have an exponential dependence on the offset amount of the high-k gate dielectric to the gate edge at a ratio of approximately 1 decade per 5nm of high-k offset.

The process flow of the P-I-N TFET is changed so that during the gate patterning, only the gate polysilicon and metal are etched and leaving the high-k dielectric intact. Then the gate dielectric is etched after a seal nitride spacer is patterned to form a protruding high-k region. This reduces the amount of dielectric damage at the gate edge during the high dose source/drain implants.

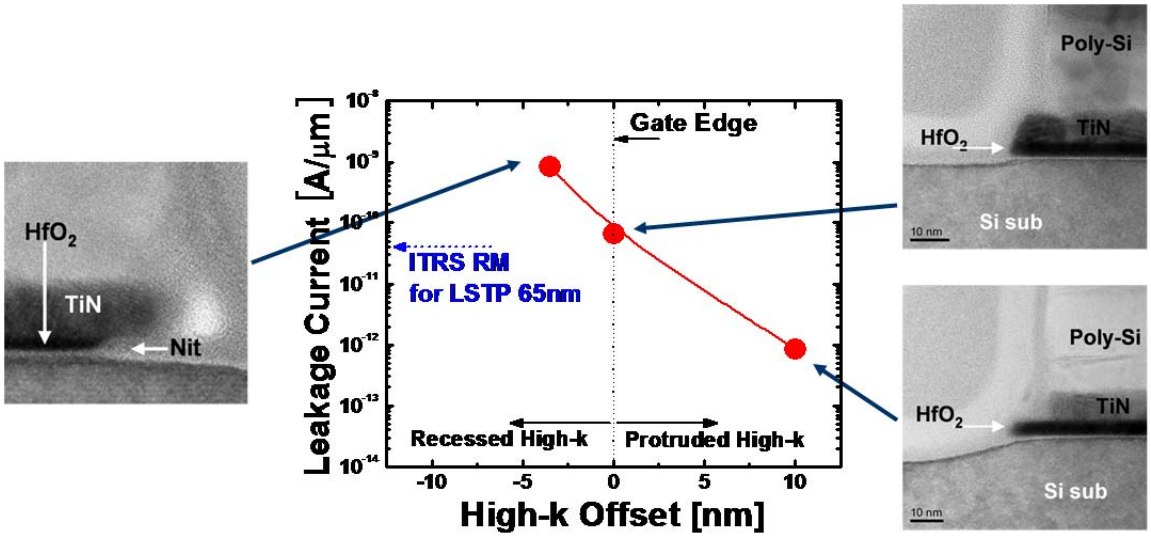
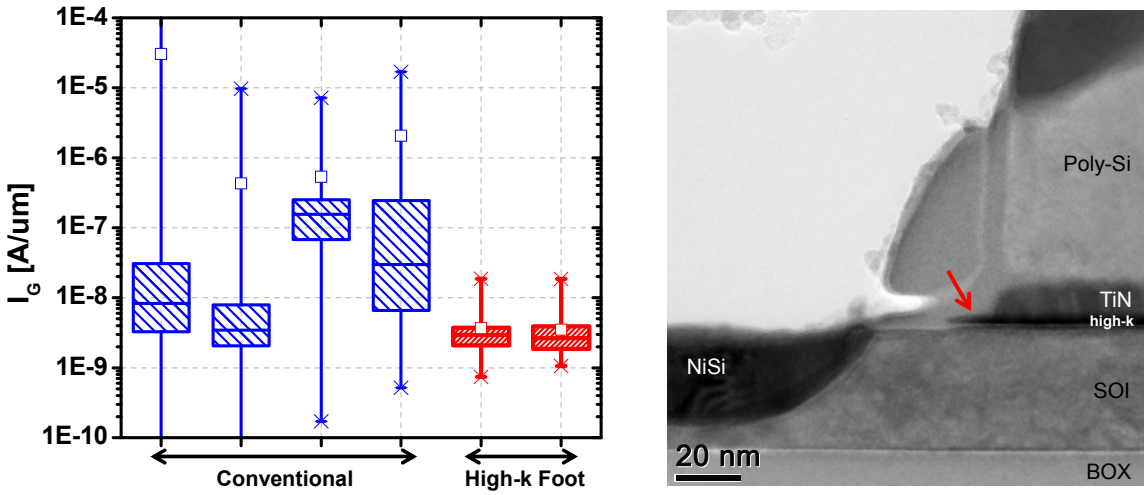


Figure 3.6 Effects of gate edge profile on OFF state leakage [3.7]

The patterned gate stack of a P-I-N TFET is shown in Figure 3.7 (b). As can be seen in the gate leakage comparison of Figure 3.7 (a), the mean gate leakage value shown as white squares was measured to decrease by at least 2 orders of magnitude. Also, the variation in the gate leakage was reduces substantially by using the high-k foot process.



(a) **(b)**

Figure 3.7 **(a)** Gate leakage current comparison between P-I-N TFET wafers with and without the protruding high-k foot process **(b)** TEM image showing the protruding high-k foot under the metal and polysilicon gate.

With the reduced gate leakage and improved gate leakage uniformity using the LOCOS isolation and modified high-k etch process this process was chosen as the baseline for the P-I-N TFET process flow.

3.3 Silicon P-I-N TFET Analysis

In this section the P-I-N measurement and analysis methods are discussed where various process splits are compared. Due to the non-self-aligned nature of the process flow of the P-I-N TFET, the first thing that needs to be checked is whether the half mask for the source and drain implants aligned correctly and formed a P-I-N diode. So, we measure the I-V characteristic between the source and drain with the gate open. Diode measurement for a 0.5µm channel length device with SOI thickness of 40nm is shown in Figure 3.8 (a). As you can see, with the N+ source grounded and P+ drain biased, a rectifying behavior can be observed and the slope of the diode I-V curve in the forward bias region is close to 60mV/dec showing a well behaved diode. If a misalignment occurs we would end up with the either N+ or P+ implanting both the source and drain and end up with an NMOS or PMOS. With this diode measurement we can verify that the half mask did not misalign and form a MOSFET accidently. We can also see the current ranges where the device will be limited by series resistance. From this measurements we see that current levels above 1µA/µm would be limited by the series resistance coming from the undoped channel and source/drain contacts. Another characteristic of this P-I-N TFET that we can obtain from the diode measurements would be the OFF state current of this device. By looking at the reverse bias measurement we can tell what OFF current levels to expect on an I_D-V_G measurement for the corresponding drain bias. Since steep subthreshold behavior in these tunneling devices usually occurs at low current ranges, especially in silicon [3.8, 3.9], we want the reverse bias current levels to be as low as possible for the intended operation voltage. Steep turn-on/turn-off behavior is generally not observed when these levels are high.

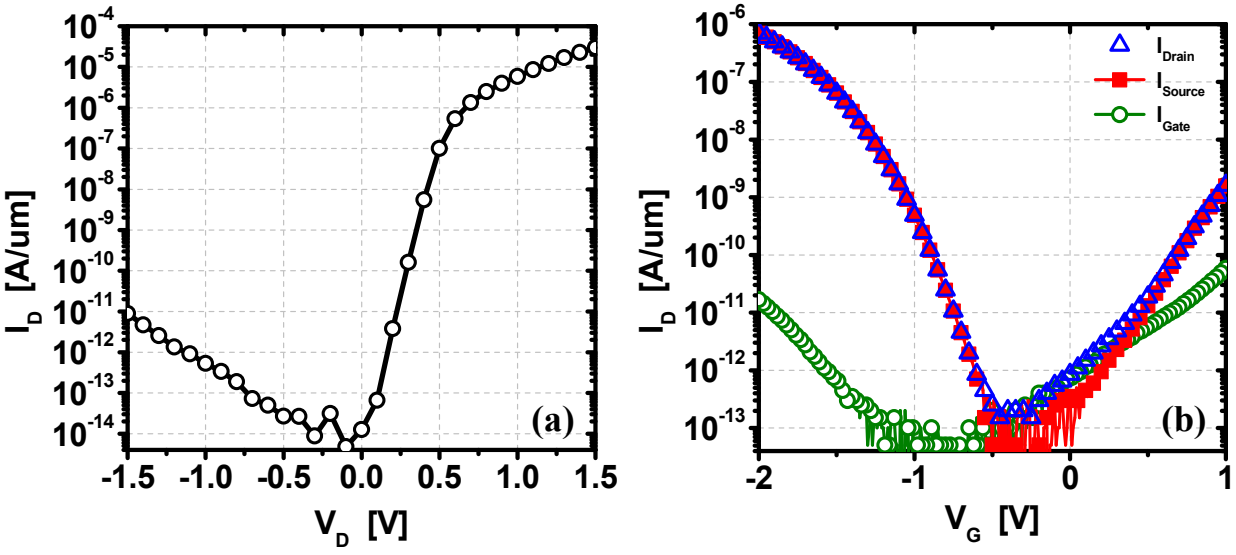


Figure 3.8 **(a)** P-I-N structure TFET diode measurement of a L_G=0.5µm device with gate open showing rectifying behavior **(b)** P-I-N TFET I_D-V_G measurement (PFET) with V_D (P+) = -1V, T_{SOI}=40nm showing SS=128mV/dec

With the confirmation of a P-I-N structure formation, we next measure the I_D-V_G characteristics to see the TFET performance. Figure 3.8 (b) shows the PFET I_D-V_G characteristics of the same device with a drain (P+) bias of -1V. The source and gate currents are also measured to check that the current being measured is band-to-band tunneling current and not gate leakage current. We can see that even though the gate leakage is fairly low, it is comparable to the drain current between -0.5V and 0V and in this range the drain current is actually coming from the gate. So when measuring the subthreshold swing we should avoid measuring in these ranges. By plotting all three currents we can tell the source of the current. When the drain current is equal to the source current we know that it is tunneling current that is being measured. We also observe that due to the different lateral straggle of the implant species, the PFET shows a better performance compared to that of the NFET.

30Å HfO_2 high-k was used as the gate dielectric with a mid-gap workfunction gate metal. We can see that with the mid-gap workfunction gate metal, the turn-on of the device occurs around -0.5V which corresponds to approximately half the bandgap of silicon. C-V measurement of the P-I-N TFET can be easily measured in one measurement due to the fact that the source and drain provide the source for holes and electrons. C-V measurements of a 20µm gate length and 10µm wide P-I-N TFET is shown in Figure 3.9. The HfO_2 dielectric used in this device shows to give an equivalent oxide thickness of ~0.9nm.

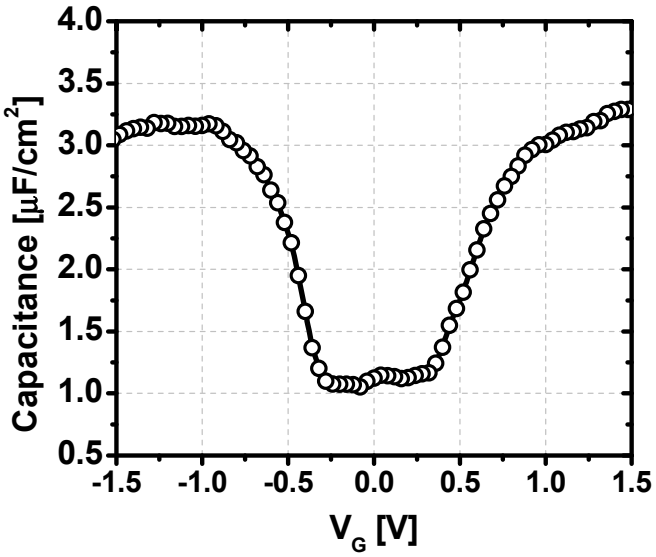


Figure 3.9 C-V measurement of a 20µm gate length and 10µm width P-I-N TFET showing an EOT of ~0.9nm

The ON current for an overdrive (V_G-V_T) of 1.5V was measured to be 7×10^{-7}A/µm. The subthreshold swing of this device does not show a sub-60mV/dec behavior and is measured to be 128mV/dec. In order to see the effect of the implant conditions, different implant conditions were included as splits in a 80nm SOI lot and the minimum subthreshold swing in each case is given in Table 3.1. All devices were annealed using spike anneal at 1050°C for 5 seconds.

Implant Condition	Slope
BF2 / 15keV / 2E15	85.2 mV/dec
BF2 / 18keV / 3e15	105.2 mV/dec
As / 55keV / 3E15	80.0 mV/dec
As / 85keV / 2E15	147.4 mV/dec

Table 3.1 BF2 and As implant conditions splits and their measured minimum subthreshold swing values 80nm thick SOI devices.

As can be seen in the above table, higher energies resulted in larger swing values due to the larger lateral straggle. And PFET devices with As as the source dopant shows a steeper swing compared to the NFET with BF_2. All the above conditions and experimental lots with a low 10keV As implant did not show sub-60mV/dec subthreshold swing. From these experimental results we can see that the abrupt lateral doping profile that is required for a sub-60mV/dec subthreshold swing shown in Figure 3.2 is difficult to achieve using conventional ion implantation and anneal techniques.

SOI thickness splits were also included and are compared in Figure 3.10. Both devices were fabricated with same implant and anneal conditions. We see that the thinner 40nm SOI devices a higher ON current and better subthreshold swing distribution. This shows that the thinner SOI devices have a better gate control of the tunneling current.

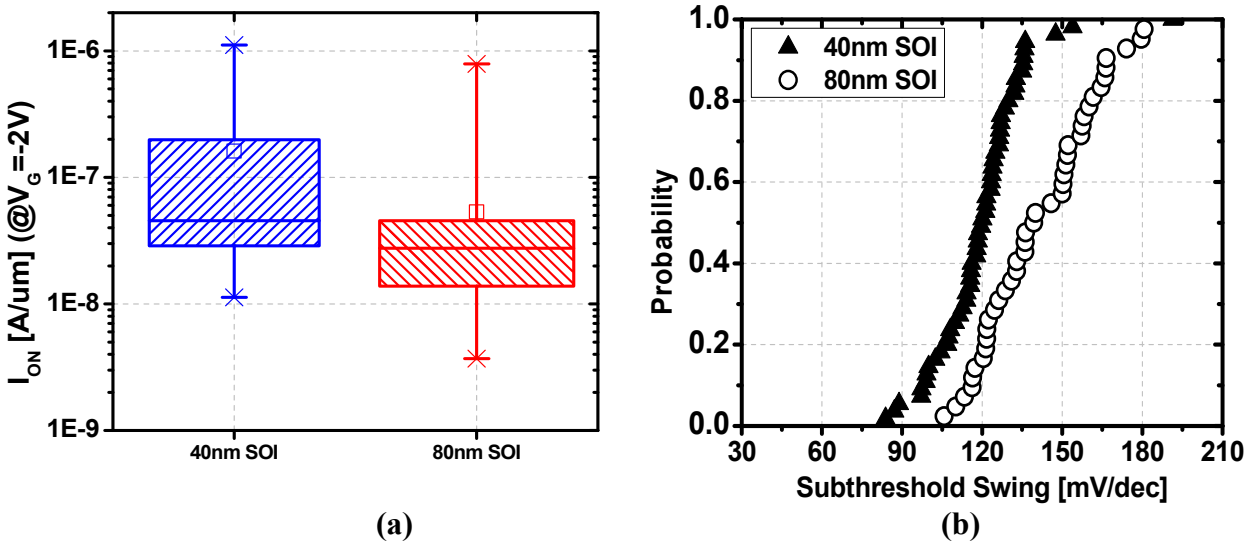


Figure 3.10 **(a)** ON current comparison of P-type P-I-N TFETs on 40nm and 80nm SOI at V_G=-2V **(b)** Subthreshold swing distribution comparison

Although the current of a TFET is generated by band-to-band generation at the source, the generated electrons or holes have to flow through the channel to the drain just like a MOSFET. So, a similar gate length dependence of the device current is expected in a P-I-N TFET. As can be seen in the following figure, the mean value of the ON current (open squares) is shown to decrease with longer gate lengths. This shows that TFET performance can be improved with physical scaling of the device dimensions.

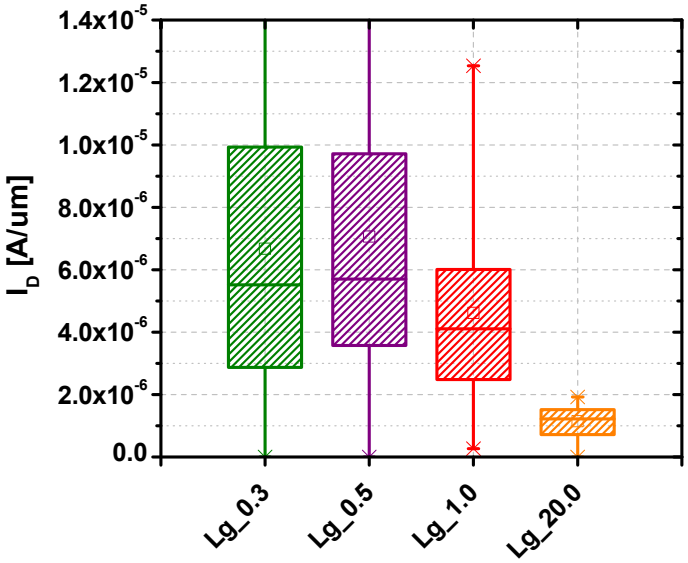


Figure 3.11 ON current versus gate length of P-type P-I-N TFETs at V_G=-2V

3.4 Summary

A baseline fabrication process flow for the silicon P-I-N structure TFET on SOI wafers was developed. Issues concerning gate stringers due to mesa isolation technique were overcome through LOCOS isolation developed for SOI wafers. The gate etch process is modified by forming a high-k foot in order to reduce the gate leakage caused by damages to the gate dielectric edge during gate etch, subsequent cleans and ion implantation process. It is shown that the gate leakage can be reduced by a considerable amount with improved uniformity by using this process.

When measuring these P-I-N structure TFETs, the diode behavior between the source and drain should always be measured to confirm the alignment of the source/drain implant half mask. All terminal current should be examined when measuring the I_D-V_G characteristics to measure real tunneling current. C-V behavior can be measured easily due to the structure having reservoirs for both electrons and holes. It is shown through simulation that a very abrupt lateral profile is required to achieve sub-60mV/dec subthreshold swing. Experimental results show the degradation in subthreshold swing with higher energy implant due to increased lateral straggle. SOI thickness dependence studies show a thinner SOI is preferred for better gate control resulting in lower subthreshold swing and higher drive current.

TFET ON current is measured to have similar gate length dependence seen in a MOSFET and will benefit from physical scaling of the device dimensions.

3.5 References

[3.1] W. M. Reddick and G. A. Amaratunga, "Silicon surface tunnel transistor," *Applied Physics Letters*, vol. 67, no. 4, pp. 494-496, July 1995

[3.2] C. Aydin, A. Zaslavsky, S. Luryi, S. Cristoloveanu, D. Mariolle, D. Fraboulet, S. Deleonibus, "Lateral interband tunneling transistor in silicon-on-insulator," *Applied Physics Letters*, vol. 84, no. 10, pp. 1780-1782, March 2004

[3.3] K. K. Bhuwalka, M. Born, M. Schindler, M. Schmidt, T. Sulima and I. Eisele, "P-channel tunnel field-effect transistors down to Sub-50 nm channel lengths", *Japanese Journal of Applied Physics*, 45, pp. 3106-3109, 2006

[3.4] W.-Y. Loh, K. Jeon, C. Y. Kang, J. Oh, P. Patel, C. Smith, J. Barnett, C. Park, T.-J. King Liu, H.-H. Tseng, P. Majhi, R. Jammy and C. Hu, "Sub-60nm Si Tunnel Field Effect Transistors with I_{ON} > 100μA/μm", *Proceedings of ESSDERC*, pp. 162-165, Sept. 2010

[3.5] W.-Y. Loh, K. Jeon, C. Y. Kang, J. Oh, T.-J. King Liu, H.-H. Tseng, W. Xiong, P. Majhi, R. Jammy, and C. Hu, "Highly scaled (L_g ~ 56nm) gate-last Si tunnel field-effect transistors with I_{ON} > 100μA/μm", *Solid-State Electronics*, vol. 65–66, pp. 22-27, Nov.–Dec. 2011,

[3.6] Taurus MEDICI User Manual, Version A-2007.12, Synopsys, Mountain View, CA

[3.7] C. Y. Kang, R. Choi, S. H. Bae, S. C. Song, M. M. Hussain, C. Yong, D. Heh, G. Bersuker, and B. H. Lee, "Effects of Optimization of Gate Edge Profile on Sub-45 nm Metal Gate High-k Dielectric Metal Oxide Semiconductor Field Effect Transistors Characteristics", *SSDM* 2006

[3.8] W. Y. Choi, B.-K. Park, J. D. Lee, and T.-J. King Liu, "Tunneling Field-Effect Transistors (TFETs) With Subthreshold Swing (SS) Less Than 60 mV/dec", *IEEE Electron Device Letters*, vol. 28, no. 8, pp. 743-745, Aug. 2007

[3.9] K. Jeon, W.-Y. Loh, P. Patel, C. Y. Kang. J. Oh, A. Bowonder, C. Park, C. S. Park, C. Smith, P. Majhi, H.-H. Tseng, R. Jammy, T.-J. King Liu, and C. Hu, "Si tunnel transistors with a novel silicided source and 46mV/dec swing," *Symposium on VLSI Technology Digest,* pp. 121-122, June 2010

Chapter 4

Achieving sub-60mV/dec Subthreshold Swing

4.1 Introduction

As shown in the previous chapter, the doping profile required to achieve sub-60mV/dec subthreshold swing in a P-I-N structure is difficult to form with the conventional ion-implantation technique. Forming a highly doped junction with a sharp lateral profile is difficult to achieve due to the lateral straggle. The subsequent heat process to activate the dopants and anneal out the defect caused by the ion bombardment worsens the sharpness of this profile. This shows that a different technique needs to be used in forming the source junction profile where the tunneling is to occur.

It has been known that dopants tend to segregate during nickel silicidation forming a highly concentration region at the silicide to silicon interface [4.1]. Figure 4.1 shows the doping profiles achieved with nickel silicide. Segregation of both boron and arsenic can be seen at the silicide to silicon interface with high concentration and abrupt profile.

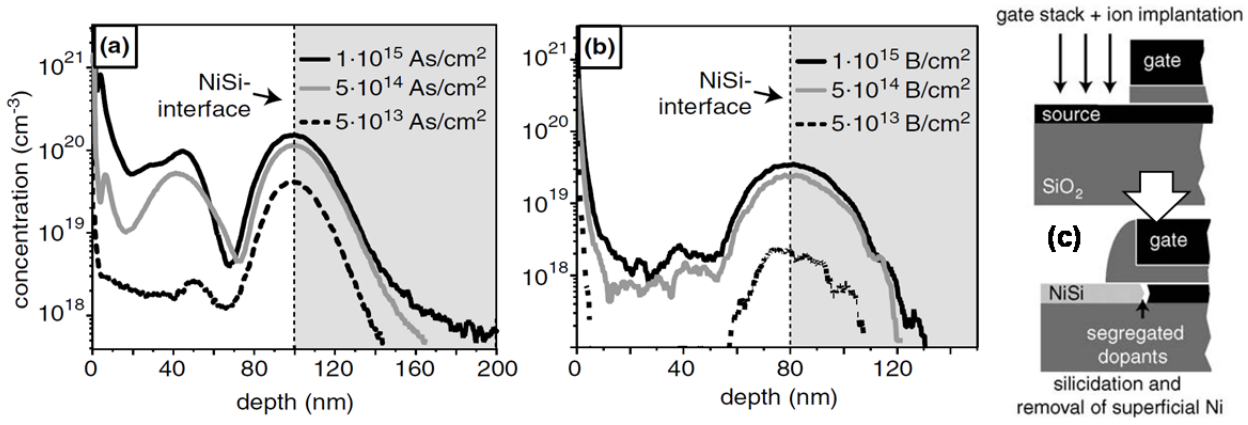


Figure 4.1 **(a)** Arsenic and **(b)** boron SIMS profiles of segregated dopant profiles at NiSi/Si interface. **(c)** Dopant segregation process flow using nickel silicide. [4.1]

In this chapter the use of dopant segregation using nickel silicide is explored using the baseline process flow developed for the silicon P-I-N structure TFET. [4.2]

4.2 Silicided Source TFET Fabrication

Since the P-I-N structure TFET can be operated both as an n-type and a p-type device, it possesses an unwanted ambipolar behavior where the drain side tunneling can cause the off state current to increase. And since dopant segregation using nickel silicidation will cause both boron and arsenic to pile up enhancing tunneling in both the source and drain, a fabrication process flow where only the source side doping profile will be improved needs to be developed. We achieve this by a recess etch only in the source side, depriving the source silicidation process of silicon and causing the silicide to encroach under the gate in the source region. The drain side where all the silicon is intact will not undergo this process leaving a less abrupt drain profile. Figure 4.2 shows the process flow of the silicided source TFET using this technique. P-type devices where chosen for fabrication due to the lower diffusion coefficient and better segregation characteristic of arsenic seen in Figure 4.1.

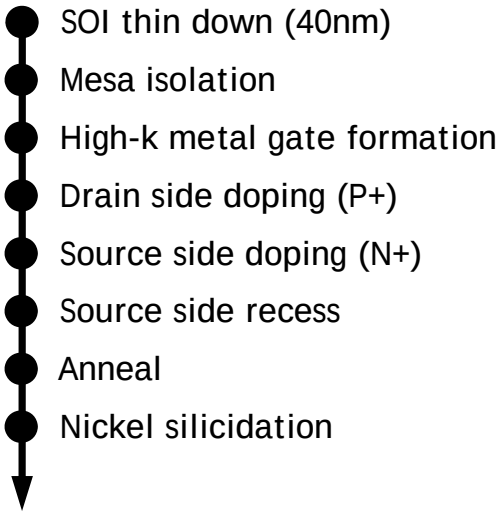


Figure 4.2 Fabrication process flow of silicided source TFET

The starting wafers used were 100nm thick SOI wafers. The silicon substrate was thinned down to 40nm using thermal oxidation in order to have the dopant junction reach the buried oxide reducing the junction area and thus the leakage current. Also, this allows the nickel to fully silicide the source after recess etch. Mesa isolation was used in order to form the active regions for a simpler process. High-k gate dielectric (HfO_2 30Å) is used to achieve low EOT and a mid-gap workfunction metal gate was used as the gate stack. The gate is patterned using a nitride hard mask. Then a thin seal nitride (~100Å) is deposited to protect the gate and gate dielectric edge during subsequent processes. The drain is formed using a photoresist half mask aligned to the center of the gate to cover the source side followed by boron ion-implantation. (BF_2, 3×10^{15} cm^{-2}, 5keV) Then the drain side is covered and source implantation (As, 3×10^{15} cm^{-2}, 5keV) using a reverse mask. The source and drain dopants were annealed using spike anneal for 5 seconds at 1070°C. Then the source region was recessed using dry etch leaving about 20nm of silicon followed by a two-step nickel silicidation. A control wafer where the source region was not recess etched was also included in the split in order to comparison.

Figure 4.3 shows the cross-section diagram and cross-section SEM images of the fabricated silicided source TFET in the source and drain region. It can be seen in the figure that the substrate thickness in the source region is about half (20nm) that of the drain region due to the source-side silicon recess etch. And due to this recess etch, the silicidation reaction in the source side was deprived of silicon causing the nickel silicide to encroach into the channel under the gate pushing the dopants in. Where in the drain region which wasn't subject to recess etch, the silicide has not overlapped with the gate. The overlap of the silicide to gate in the source end was measure to be a few 100nm which caused short channel length (L_G < 0.25µm) devices to show an electrical short between the source and drain.

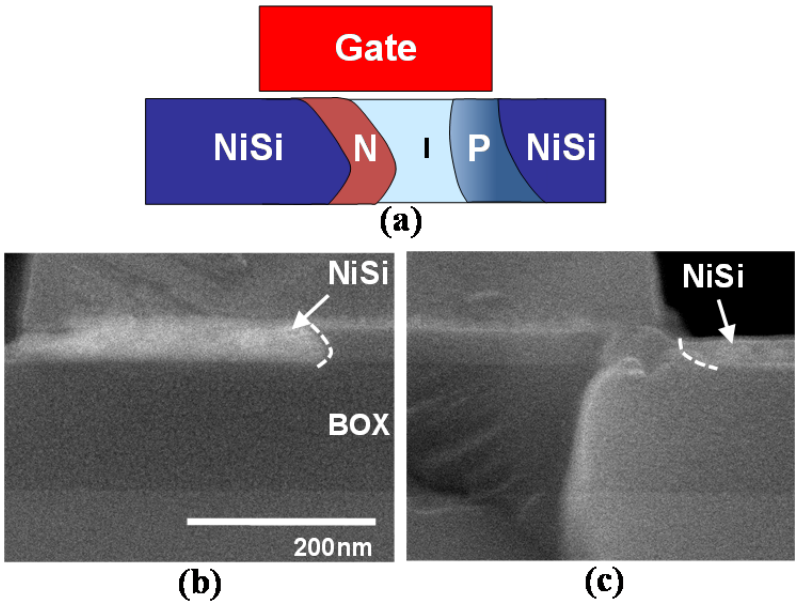


Figure 4.3 **(a)** Silicided source TFET structure. Cross-section SEM images of silicided source TFET at **(b)** source (N+) region showing silicide encroachment under gate and **(c)** drain (P+) region without any gate-to-silicide overlap

Cross-section SEM images of the control wafer were also taken and are shown and compared with the silicided source TFET in Figure 4.4. It can be seen that without the source-side recess etch, the silicide did not encroach under the gate.

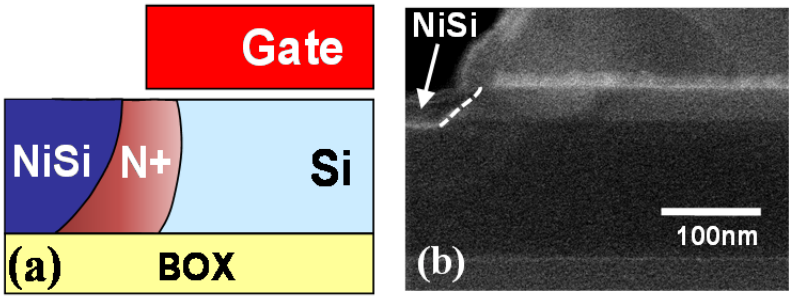


Figure 4.4 Control P-I-N TFET **(a)** Structure and **(b)** cross-section SEM image showing no silicide to gate overlap.

4.3 Analysis of the Silicided Source TFET

Electric characteristics of the silicided source TFET were measured and the I_D-V_G characteristics of a 20µm channel length device is shown in Figure 4.5 where the drain bias was held at -1V. As can be seen from the figure, a sub-60mV/dec subthreshold swing is observed and averages 46mV/dec over a decade of drain current at the steepest region. The subthreshold swing is not constant as expected from a TFET and gradually increases with saturation of tunneling occurring around 0.1µA/µm. The leakage current of the device is measured to be very low in the 10^{-14}A/µm range due to the long channel length and an I_{ON}/I_{OFF} ratio of 7×10^7 is achieved for a 1V operation which is the highest reported for silicon TFETs. Table 4.1 compares the silicided source with reported TFET data.

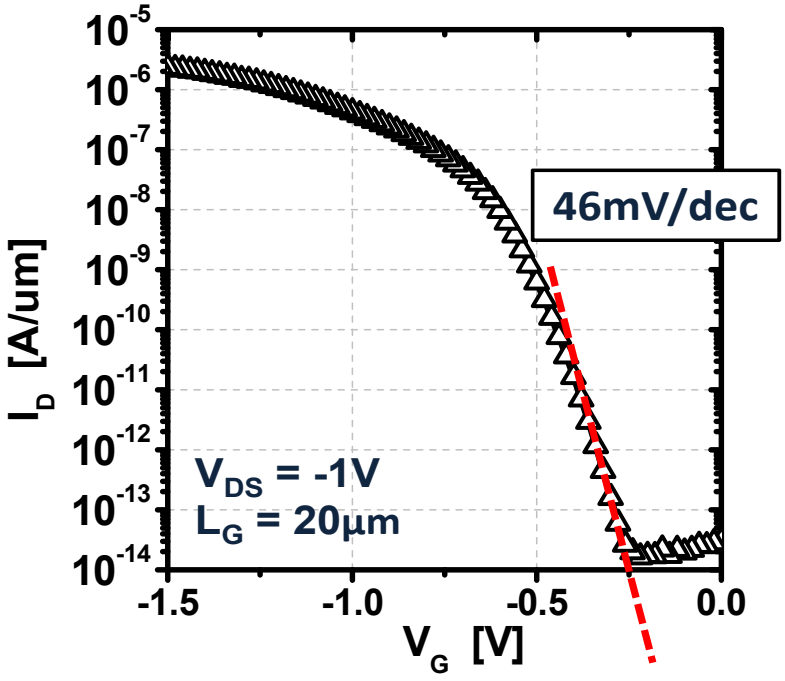


Figure 4.5 Measured I_D-V_G of silicided source TFET (PFET)

	Ref. [4.3]	Ref. [4.4]	Ref. [4.5]	This Work
SS (mV/dec)	52.8	42	~300	46
I_{ON} (µA/µm)	12.1	0.01	1E-4	1.2
I_{ON}/I_{OFF}	1E4	1E4	1E2	7E7

Table 4.1 Comparison of silicided source TFET to reported silicon TFETs. (V_{DS}=V_{GS}-V_{BTBT}=1.0V)

The silicided source TFET is compared to the control device processed with the P-I-N baseline flow without the source side recess etch and is given in Figure 4.6. Since the control device was not subject to dopant segregation with the nickel silicide and has a more gradual source doping profile, it shows a very slow increase in the drain current with respect to the gate bias and a lower drive current compared to the silicided source TFET. And as can be seen in the subthreshold swing versus drain current plot, not only does the control device not show a subthreshold swing that is lower than 60mV/dec but the swing value increases much faster than the silicided source TFET with respect to the drain current. In the silicided source TFET we can see that the sub-60mV/dec swing region extends over almost 3 decades of current and shows a very gradual increase. This shows that the silicide source TFET has a good control of band-to-band tunneling by the gate.

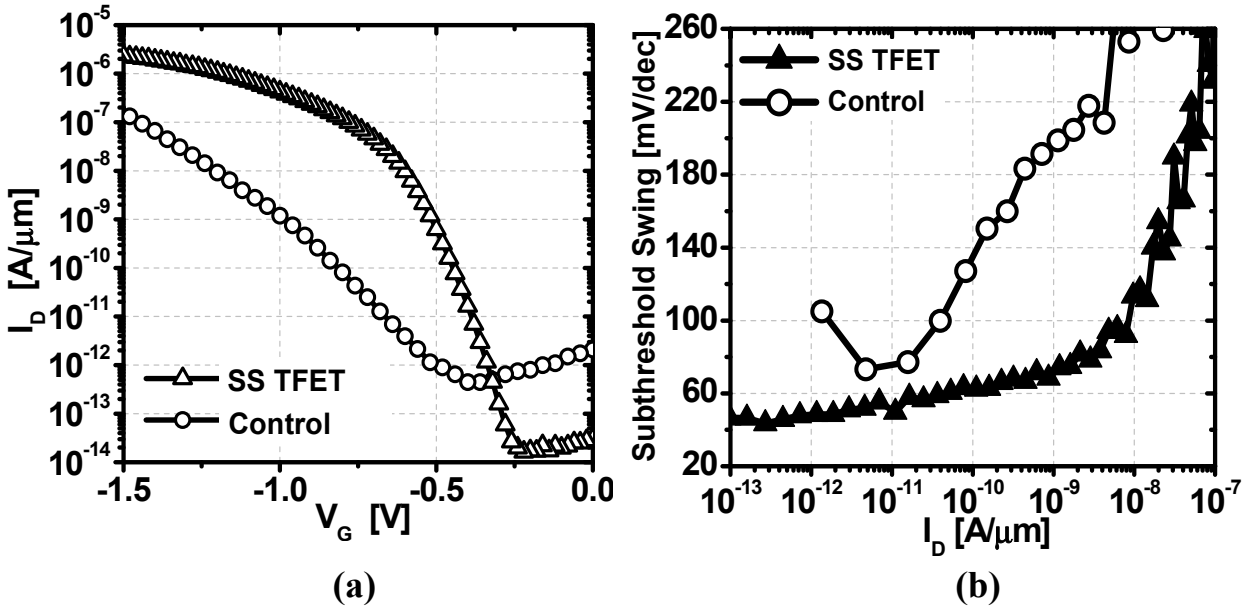


Figure 4.6 **(a)** I_D-V_G and **(b)** Subthreshold swing vs. drain current comparison between the silicided source (SS) TFET and control TFET

Statistical distribution of the minimum subthreshold swing of the silicided source TFET were also analyzed and is shown in Figure 4.7. The silicided source TFET shows more than 30% to have a sub-60mV/dec subthreshold swing whereas the control wafer does not show any. The variation in the subthreshold swing for the silicide source TFET is thought to be caused by the variation in the silicide edge roughness, shape and segregated dopant profile. But in comparison to the control TFET, the silicided source TFET shows a tighter distribution in the subthreshold swing indicating a better process control can be achieved with the dopant segregation compared to implant and anneal alone. The silicided source TFET was successfully reproduced in a different lot within a two month time frame achieving 47mV/dec subthreshold swing with similar distribution shown in Figure 4.7.

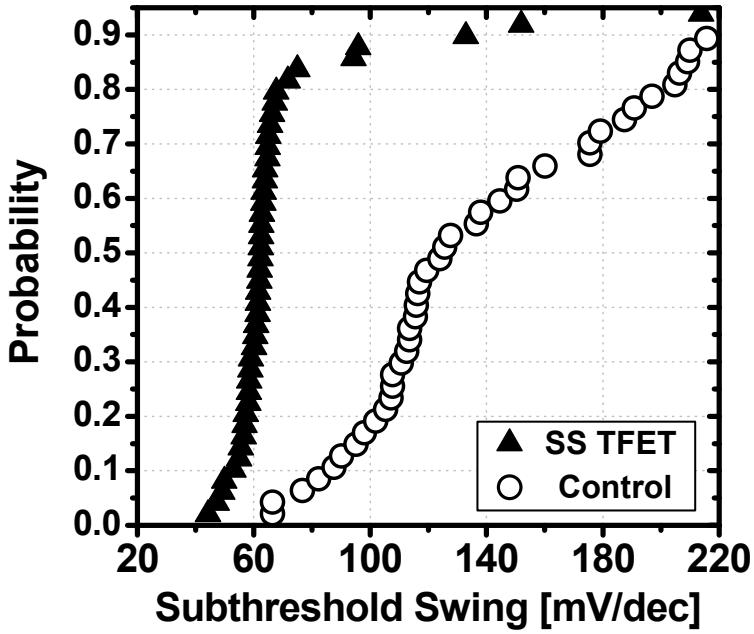


Figure 4.7 Subthreshold swing distribution comparison between silicide source TFET and control TFET

The silicided source TFET was also fabricated using in-situ steam generation (ISSG) SiO_2 as the gate dielectric in place of HfO_2. Figure 4.8 shows the C-V measurement and statistical distribution of the subthreshold swing for both gate dielectrics. This shows the thinner equivalent oxide thickness (EOT) of the high-k dielectric with better gate control results in improved subthreshold swing.

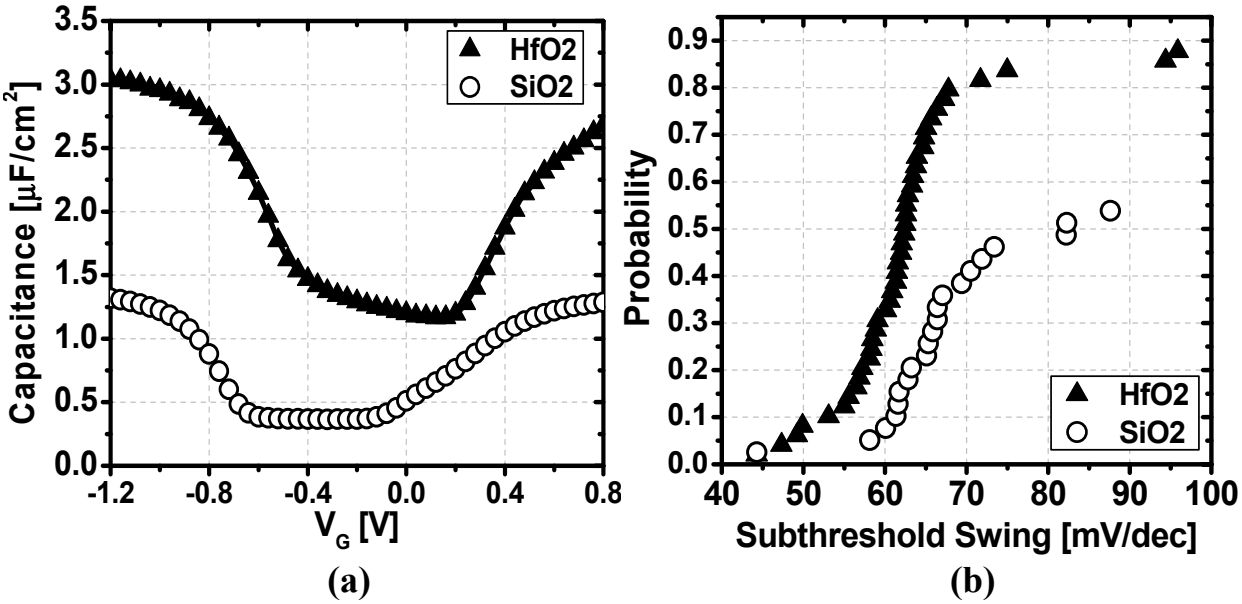


Figure 4.8 **(a)** C-V and **(b)** Subthreshold swing distribution comparison between HfO_2 and SiO_2 gate dielectric

I_D-V_D characteristics of the silicided source TFET is shown in Figure 4.9 (a). We see that the drain current exhibits a nonlinear relation to the drain bias at low voltages. This is expected from a P-I-N structure TFET since the drain field has some control of the tunneling behavior at low voltages which has an exponential characteristic. I-V measurements with just biasing the source and drain of the device shows a rectifying behavior (Figure 4.9 (b)) confirm the formation of a diode between the source and drain. This shows that the device cannot be a MOSFET.

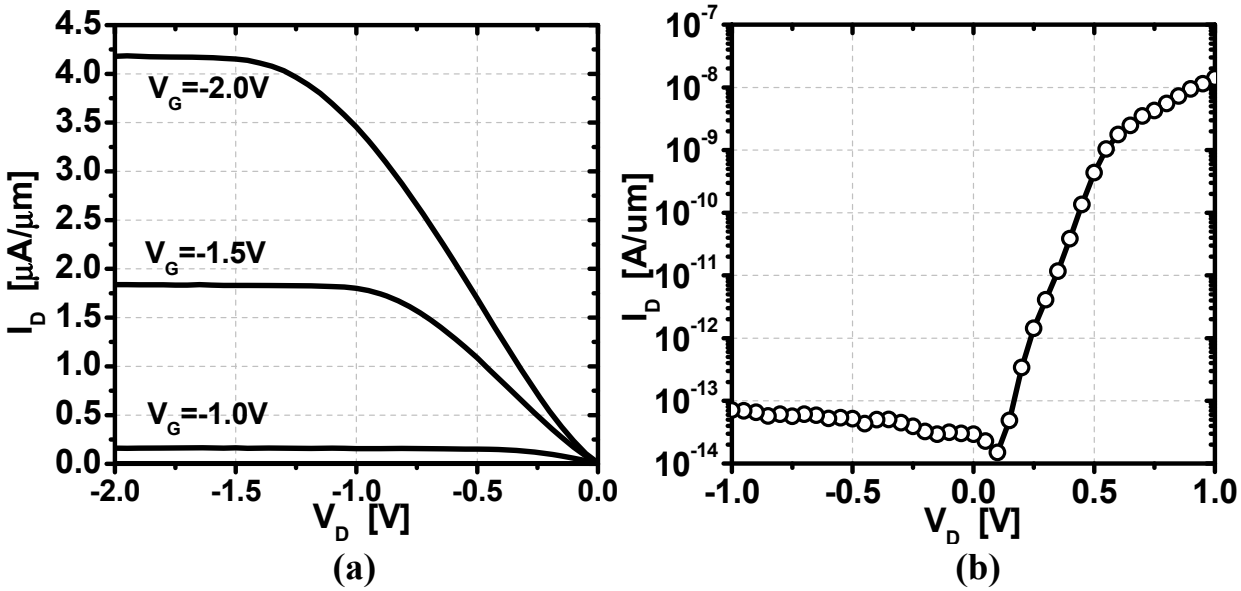


Figure 4.9 (a) I_D-V_D and **(b)** Diode characteristics of silicided source TFET.

Although a distinct difference in the characteristics between the silicided source TFET and control TFET can be seen, we need to confirm the existence of segregated arsenic dopants at the silicide to silicon channel interface and that the nickel silicide did not pass the dopant front. In order to do this temperature dependence measurements of the inherent diode in the silicided source TFET were conducted and the barrier height for holes and electrons to the nickel silicide were calculated. The Schottky diode equation which is written as

$$I = AA^{**}T^2 exp\left(-\frac{q\phi_B}{kT}\right)\left[exp\left(\frac{qV}{kT}\right) - 1\right] \tag{4.1}$$

where A is the area of the diode, A^{**} the effective Richardson constant and ϕ_B the barrier height, can be manipulated to the following equations 4.2 and 4.3 for forward and reverse bias respectively [4.6].

$$ln\left(\frac{I_F}{T^2}\right) = \ln(AA^{**}) - \frac{q(\phi_B - V)}{kT} \tag{4.2}$$

$$ln\left(-\frac{I_R}{T^2}\right) = \ln(AA^{**}) - \frac{q\phi_B}{kT} \tag{4.3}$$

By plotting $ln(I/T^2)$ against $1/T$ and finding the slop we can calculate the barrier height for electrons and holes. Figure 4.10 (a) and (b) shows the measurement results for the forward bias and reverse bias respectively. As can be seen in Figure 4.10 (a), the electron barrier height was measured to be relatively constant at ~0.7eV for a forward bias range of 0.1V to 0.5V which agrees with known values for NiSi. However, in the reverse bias case, the hole barrier height was extracted to be 0.77eV for a low reverse bias of -0.1V which is higher than the reported value of ~0.4eV for NiSi. [4.7] The hole barrier height is also found to decrease with larger reverse bias. This indicates the presence of an n-type dopant layer at the silicide interface as illustrated in Figure 4.11.

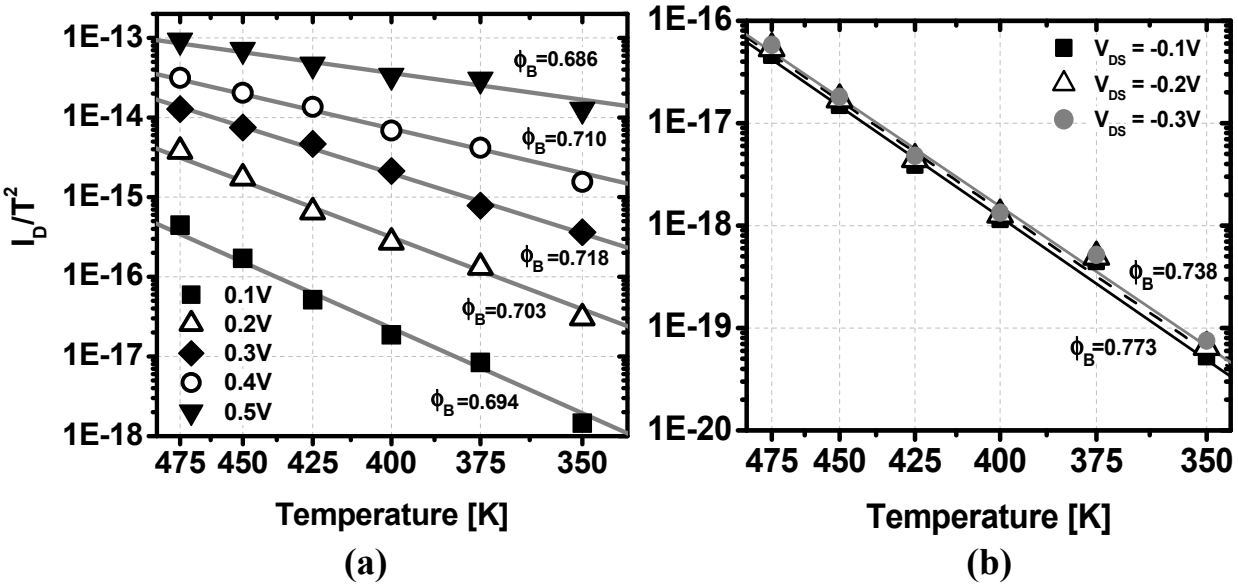


Figure 4.10 Temperature dependent diode measurement **(a)** Forward bias and **(b)** Reverse bias measurement with calculated electron and hole barrier height respectively.

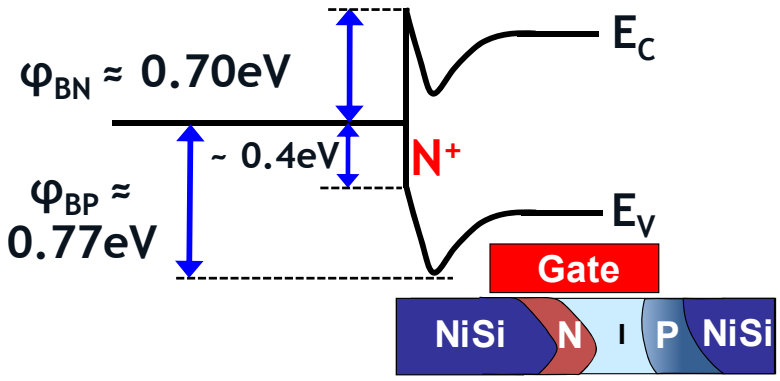


Figure 4.11 Energy band diagram of silicided source TFET at silicide to source interface showing band bending due to pile up of N-type dopants.

In addition to the sharp profile caused by dopant segregation, the wedge shape formed between the silicide and the gate dielectric plays a critical role in this device. It confines the band-bending to the thin wedge-shaped pocket region, thus creating a high electric field especially in the thin corner of the wedge. It also changes the direction of tunneling to be more vertical than lateral which gives the gate a better control of the turn-on. This is confirmed through MEDICI simulations [4.8]. As can be seen in Figure 4.12, simulation results for the silicide geometry that creates a thin doped pocket region shows steep turn-on matching the measured transfer characteristics, whereas simulation results for a vertical silicide interface (without the pocket) do not show sub-60mV/dec subthreshold swing. This shows that field concentration in the highly doped wedge pocket is critical in the performance of this device. The deviation between simulated and measured results seen for high current levels can be explained quantitatively by the high series resistance in the experimental device caused by the thinness of the nickel silicide (20nm) at the source. Lower series resistance and more controllable pocket geometry should improve the drive current of the device to hundreds of μA/μm [4.9].

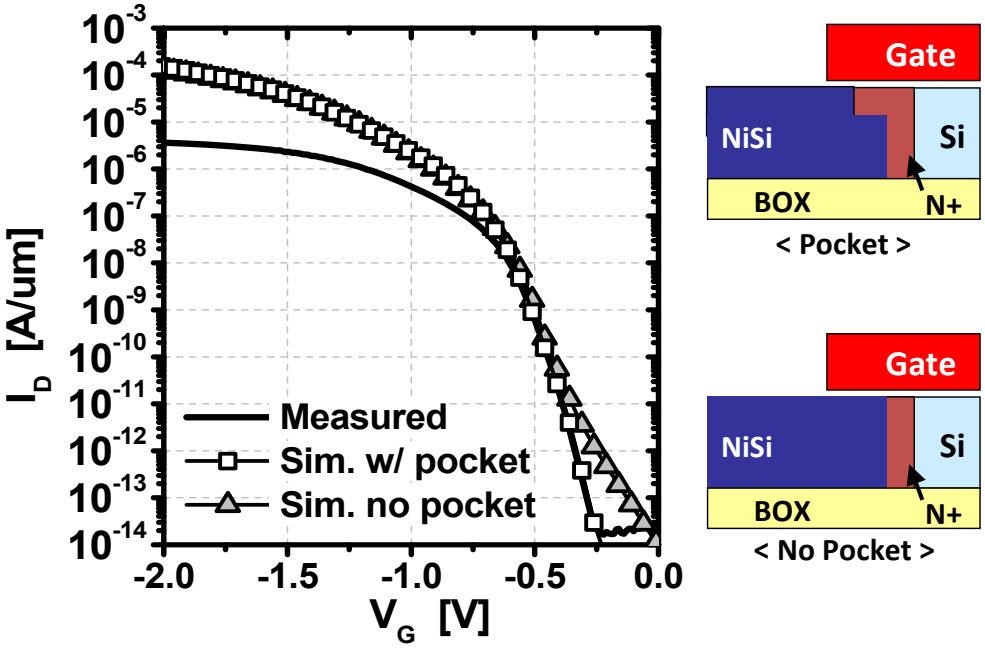


Figure 4.12 MEDICI simulation results showing the effect of the wedge shaped silicide pocket. Simulation results of structure with the pocket matches the measured current characteristics of the silicide source TFET

The effect of the wedge shape is also confirmed by measuring the gate induced tunneling currents in the source and drain separately using gate induced drain and source leakage (GIDL, GISL) measurement configurations. As can be seen in Figure 4.13, GISL current shows a sharp turn on due to the wedge shape of the silicide confining the gate electric field to a thin region at the surface. This allows for a more efficient band-to-band tunneling to occur. On the other hand, the drain side where the boron profile is more gradual and does not have the wedge shape silicide pocket, the gate field is spread out deeper into the substrate and less tunneling occurs due to the increased barrier thickness. This is displayed in the GIDL curve. The current shows a very gradual increase with the current level not reaching as high as GISL.

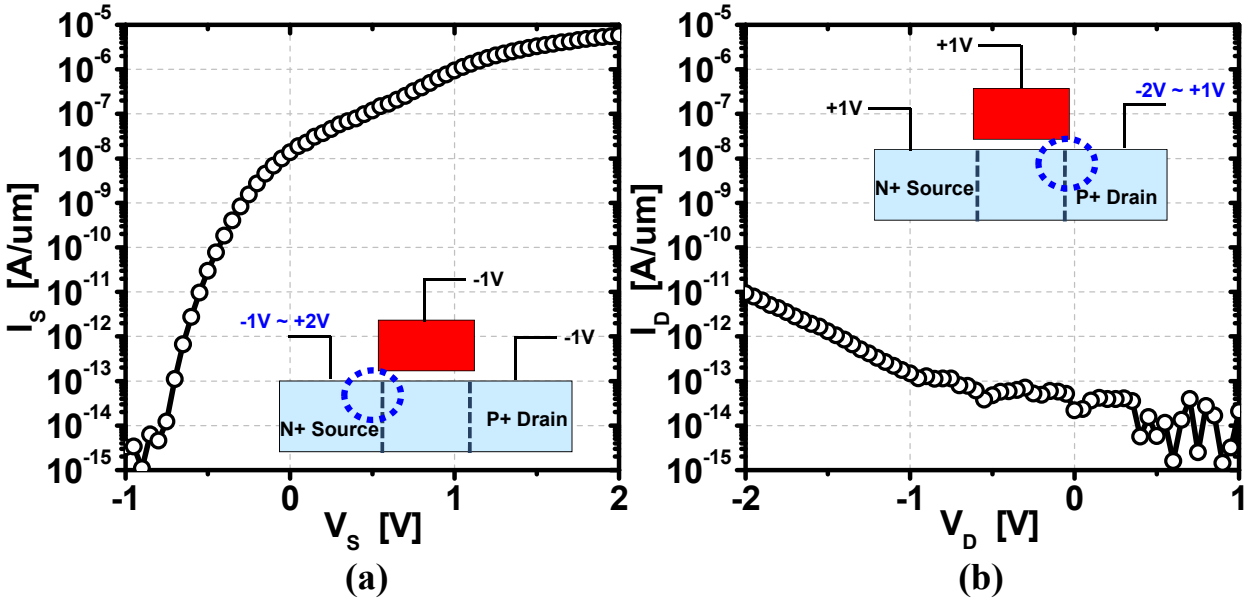


Figure 4.13 Measured results and configuration for **(a)** GISL and **(b)** GIDL showing higher GISL current with sharper turn-on.

Temperature dependent I_D-V_G measurements were carried out for the silicided source TFET to observe the change in the subthreshold swing with respect to temperature and are shown in Figure 4.14.

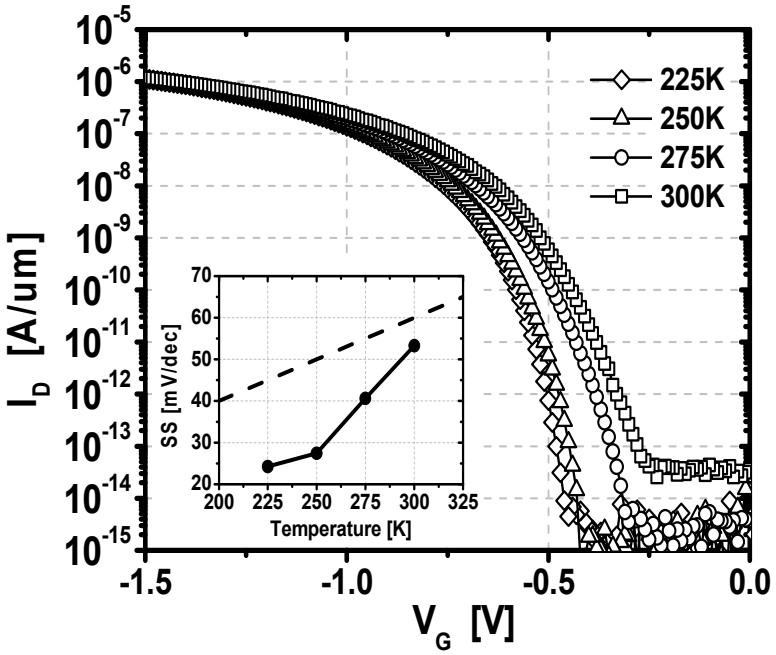


Figure 4.14 Measured I_D-V_G of silicided source TFET at various temperatures from 225K to 300K. Inset shows the measured subthreshold swing plotted against temperature.

The general belief is that band-to-band tunneling devices will not show any dependence of the current on temperature. But this is not true since the band-to-band tunneling involves tunneling between filled states in the valence band and empty states in the conduction band and the filling of these states will behave differently with respect to changes in temperature. The Fermi level which also has a temperature dependence will also contribute to the dependence of the tunneling current as well. For example, depending on the doping level and the profile of the source, the band-to-band tunneling can occur in different regions in the source. Figure 4.15 shows MEDICI device simulation results of two cases where differences in doping profiles of the source causing different temperature dependence of the tunneling characteristics. Figure 4.15 (a) and (b) shows that when the tunneling occurs within a very highly doped region, the tunneling current does not show any voltage shift due to changes in temperature. This is due to the fact that the Fermi level of a highly doped region has a weak temperature dependence characteristic. On the other hand, when the tunneling occurs in a relatively low doped region as shown in Figure 4.15 (c) and (d), the tunneling I-V curve shows a larger shift in turn-on voltage with respect to temperature since the Fermi level shows a larger temperature relation at lower doping concentrations. (Figure 4.16)

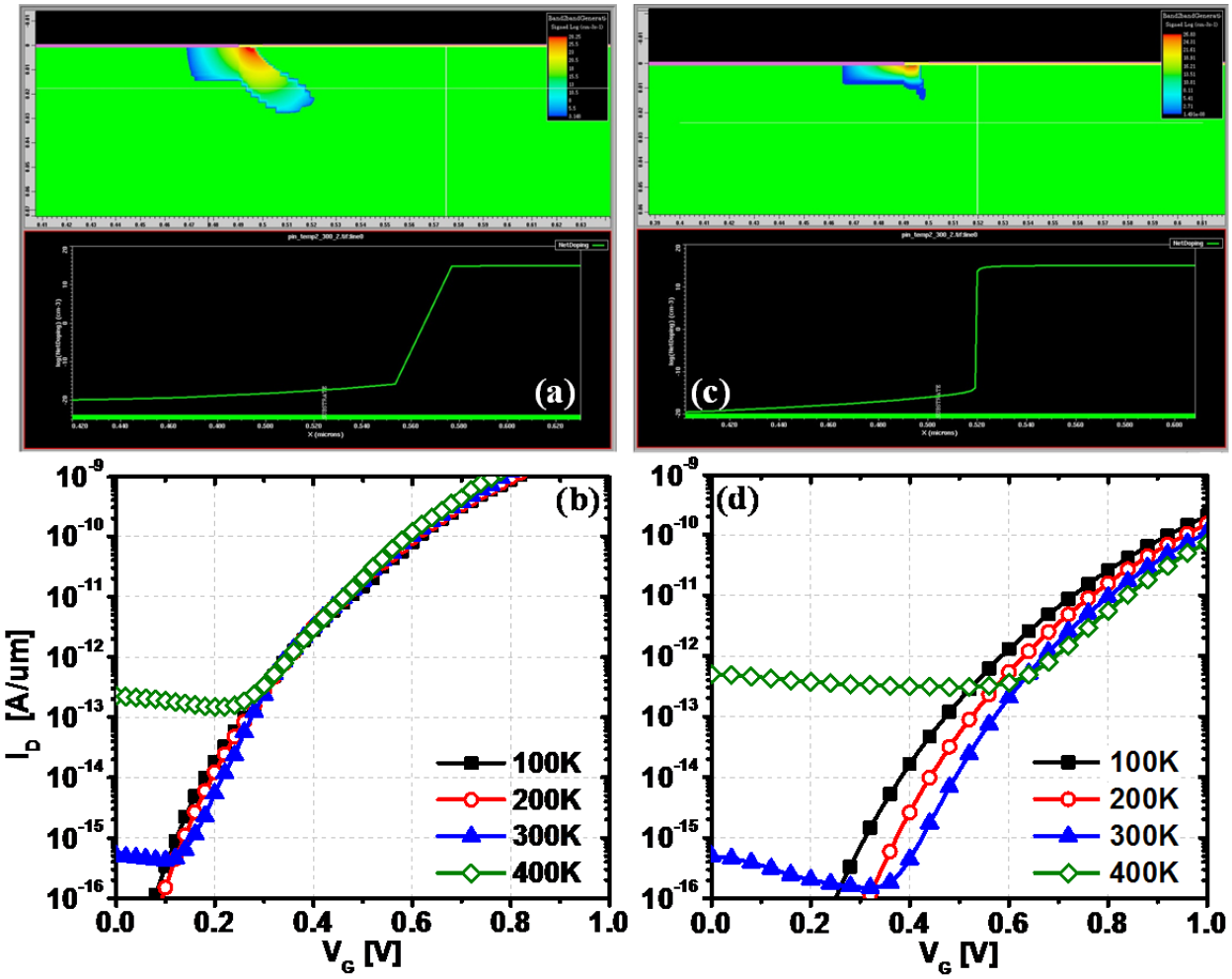


Figure 4.15 Measured results and configuration for **(a)** GISL and **(b)** GIDL showing higher GISL current with sharper turn-on.

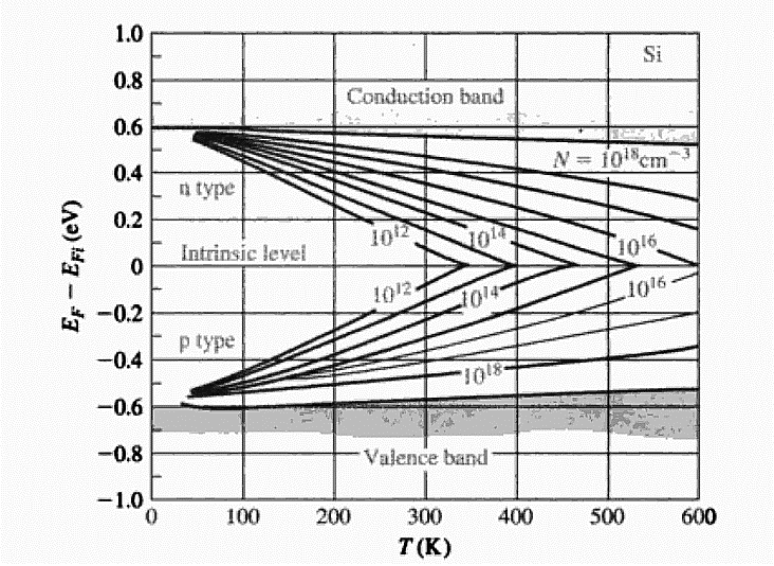


Figure 4.16 Position of Fermi level as a function of temperature for various doping concentrations [4.6]

Also, it has been shown that tunneling devices possess a different temperature dependence characteristic compared to that of a MOSFET. The tunneling device should show a nonlinear change in the subthreshold swing with respect to temperature whereas in a MOSFET it would be linear [4.10~4.11]. And we observe this dependence in the silicide source TFET. The subthreshold swing decreases with lower temperature but in a nonlinear fashion as can be seen in the inset of Figure 4.14.

In an attempt to find the optimal thickness of the SOI and silicidation conditions which would generate the dopant segregation and the wedge shape, P-I-N silicon TFETs were fabricated on SOI wafers with two different substrate thickness and nickel silicide splits varying the deposited nickel thickness and anneal temperature. Table 4.2 gives the split table for this lot.

Wafer #	Description
1	40nm SOI, 10nm Nickel (300C)
2	40nm SOI, 15nm Nickel (300C)
3	40nm SOI, 15nm Nickel (450C)
4	30nm SOI, 10nm Nickel (300C)
5	30nm SOI, 15nm Nickel (300C)
6	30nm SOI, 15nm Nickel (450C)

Table 4.2 Split table for silicided source experiment varying SOI thickness, deposited nickel thickness and silicidation temperature

Figure 4.17 (a) and (b) shows the measured PFET I_D-V_G of 20µm channel length devices for the wafer #03 and #04 listed in Table 4.2.

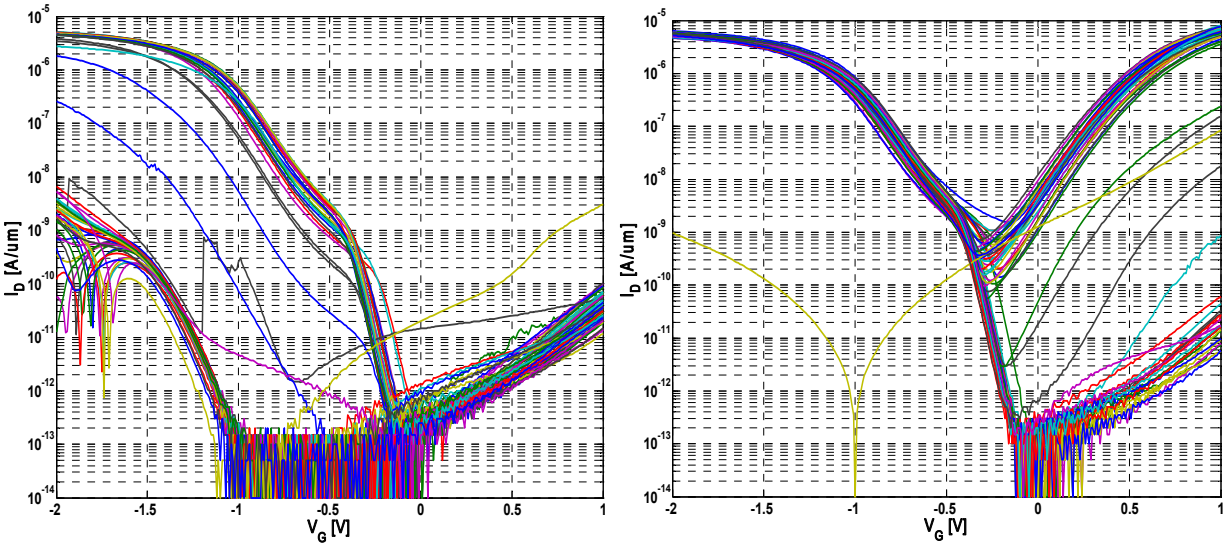


Figure 4.17 Measured I_D-V_G of L_G=20µm TFETs **(a)** Wafer #03 **(b)** Wafer #04

As can be seen from the two figures, we can see that half of the devices in wafer #03 show conventional P-I-N silicon TFET characteristics without dopant segregation and half of the devices showing to behave more like the silicided source TFET. As for wafer #04, we can see half behaving like the silicided source TFET and the other half having ambipolar behavior. This suggests that wafer #04 with the thinner SOI thickness has more encroachment of the nickel silicide and the half that is showing the ambipolar behavior has both the arsenic and boron front to be pushed in by the silicidation process.

			72.3	106.6	109.9			
		214.6	59.6	116.3	89.3	62.9		
	157.9	103.3	109.6	107.3	65.8	118.5	62.2	
104.0	113.3	109.2	105.8	100.7	61.8	62.0	110.9	62.0
110.9	110.5	87.6	65.5	60.2	61.3	61.1	60.3	62.2
103.2	109.6	110.3	55.4	62.8	61.9	64.8	61.2	61.3
	105.3	105.9	125.1	62.0	66.4	62.4	64.8	
		82.4	93.0	124.4	134.0	65.3		
			98.0	103.0	66.5			

Wafer 03

			X	62.7	167.0			
		56.1	60.3	59.5	64.4	59.5		
	58.1	61.8	61.3	103.3	97.7	95.4	66.2	
121.5	61.2	61.1	57.5	61.4	87.6	133.9	89.2	62.6
60.1	60.0	51.2	60.2	176.8	138.9	123.2	107.8	102.0
59.8	54.2	58.6	63.1	155.1	196.7	136.2	129.3	90.9
	60.2	129.0	162.4	169.0	111.6	201.2	198.1	
		103.3	201.3	194.0	187.5	205.6		
			209.1	204.2	202.5			

Wafer 04

Figure 4.18 Wafer map of measured subthreshold swing

Figure 4.18 shows the wafer map of the steepest subthreshold swing of the 20µm channel length devices measure on the wafer. We can see that the lower right half of wafer #03 and upper left half of wafer #04 show a subthreshold swing value close to 60mV/dec. These devices correspond to the devices that show and I_D-V_G close to the silicided source TFET. A thickness variation of the SOI thickness was measured in both wafers and for wafer #03, the close to 60mV/dec subthreshold swing were measured in the area where the SOI thickness was on the thinner side and as for wafer #04, the thicker side showed this behavior with the thinner lower right half showing the ambipolar behavior. This suggests that with the silicide conditions for wafer #03, the thinner SOI gives the dopant segregation in the arsenic side where the boron dopants were not segregated with the nickel silicide. This is confirmed from MOSFET measurements on the wafer.

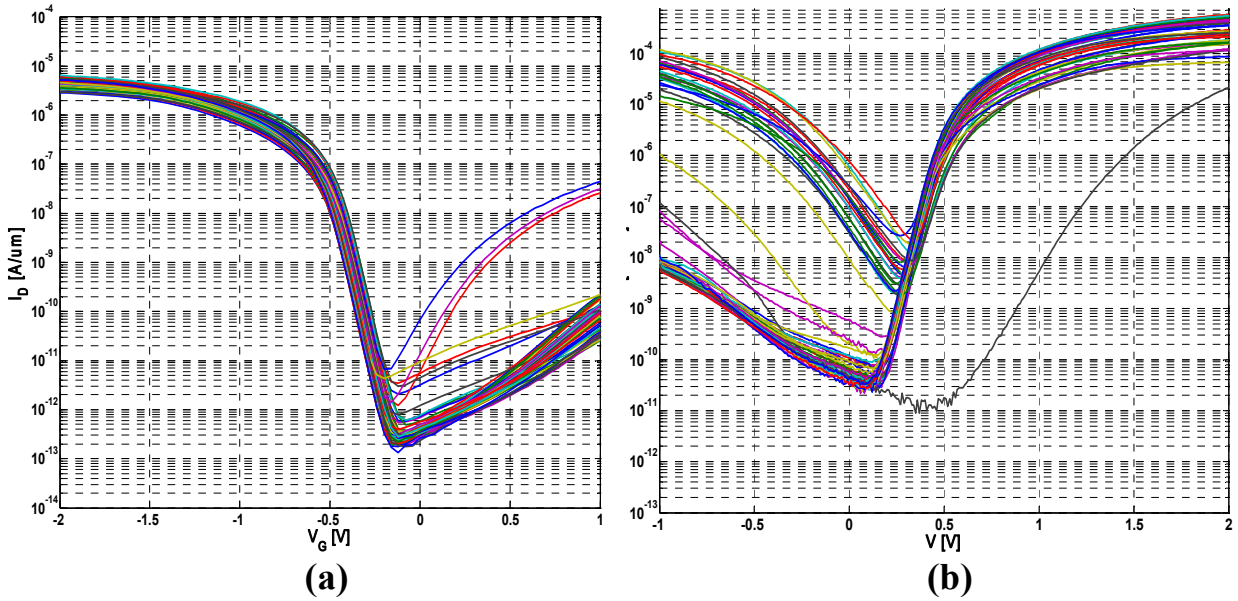


Figure 4.19 Measured I_D-V_G of MOSFETs on Wafer #03 **(a)** PMOS **(b)** NMOS

Figure 4.19 (a) and (b) shows the MOSFET ID-VG measurements of 0.6µm channel length NMOS devices and 10µm channel length PMOS devices. NMOS devices in the dies where the P-I-N TFET showed the dopant segregated like behavior shows and ambipolar behavior suggesting the nickel silicide has encroached close to or even passed the arsenic dopant front. As for the boron side, PMOS measurements show that most of the devices do not show this large ambipolar behavior which shows the boron dopant front is well beyond the silicide to silicon interface. On the other hand PMOS measurements of wafer #04 shown in Figure 4.20 shows to have a large portion of the devices showing ambipolar behavior and these dies correspond to the dies where the P-I-N TFET were measured to have the ambipolar behavior. This shows that the lower right region of the wafer with the thinner SOI thickness is sufficient to allow dopant segregation for both arsenic and boron for the silicide conditions of wafer #04.

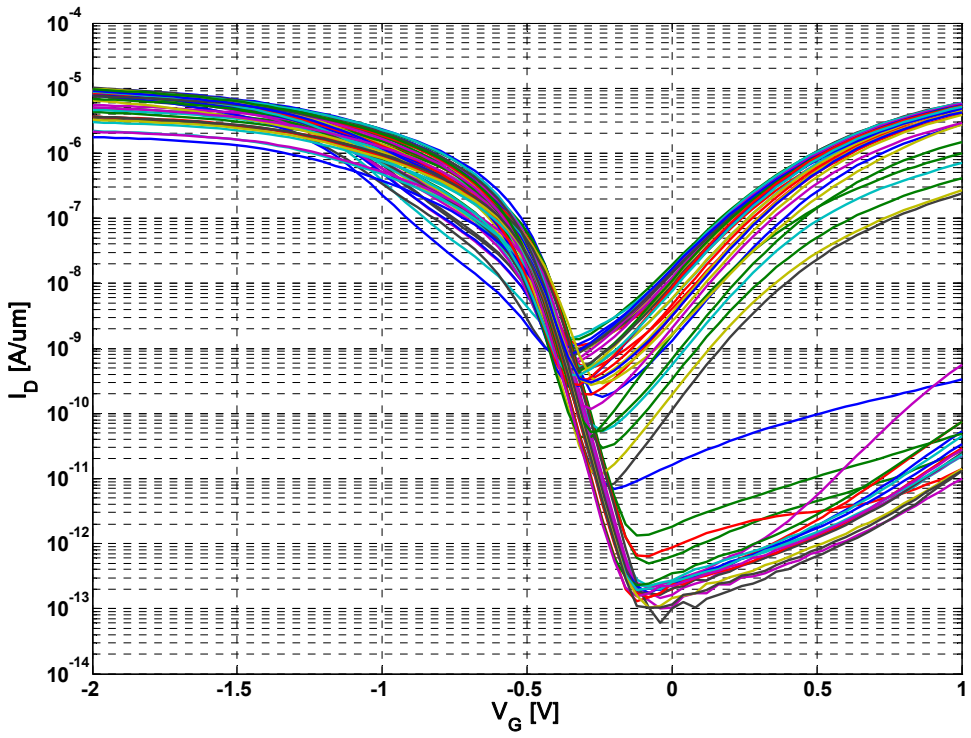


Figure 4.20 Measured I_D-V_G of PMOSFETs on Wafer #04

But as can be seen in Figure 4.18 of the wafer map of the subthreshold swing values, only a few devices showed a subthreshold swing that is less than 60mV/dec. And even in those devices that did show a sub-60mV/dec subthreshold swing, the values were very close to 60mV/dec. This suggests that although the SOI thickness of wafer #03 and #04 were sufficient enough to allow for the nickel silicide to be close to the dopant front, having a uniform SOI thickness were not helpful in forming the wedge shape of the silicide that is required for the high electric field. The uniform thickness of the source, channel and drain only allowed the nickel silicide to encroach laterally. In order to form the wedge shape of the nickel silicide, only the source region needs to be recessed with the channel region to be sufficiently thicker. This will allow the nickel silicide to form from below the channel surface and grow upward which helps in forming the wedge shape that is required for the confinement of the electric field.

4.4 Subthreshold Swing Data Quality Analysis

Steep subthreshold swing in TFETs has so far been reported only at very low drain current levels, which may be comparable to the gate leakage and/or junction leakage currents. Therefore, the quality of the data needs to be carefully analyzed. We need to show that the current which is showing the sub-60mV/dec behavior is actually coming from the band-to-band tunneling current and not from gate leakage or other sources. Also, transient effects due to charging and discharging of traps should be considered in evaluating the subthreshold swing. In this section we discuss about the analyzing quality of the measured tunneling current and subthreshold swing and propose screening criteria to eliminate any spurious data that are measured.

When gate leakage and/or junction leakage components are present, the drain current can undergo a sign change, especially at the low current levels where the steep subthreshold swing is measured. Any zero crossing plotted in a semi-log I_D-V_G plot can seem to have a very steep subthreshold swing but in reality the steepness is coming from the sign change. Figure 4.21 shows this case there the measured drain current undergoes a sign change (zero crossing).

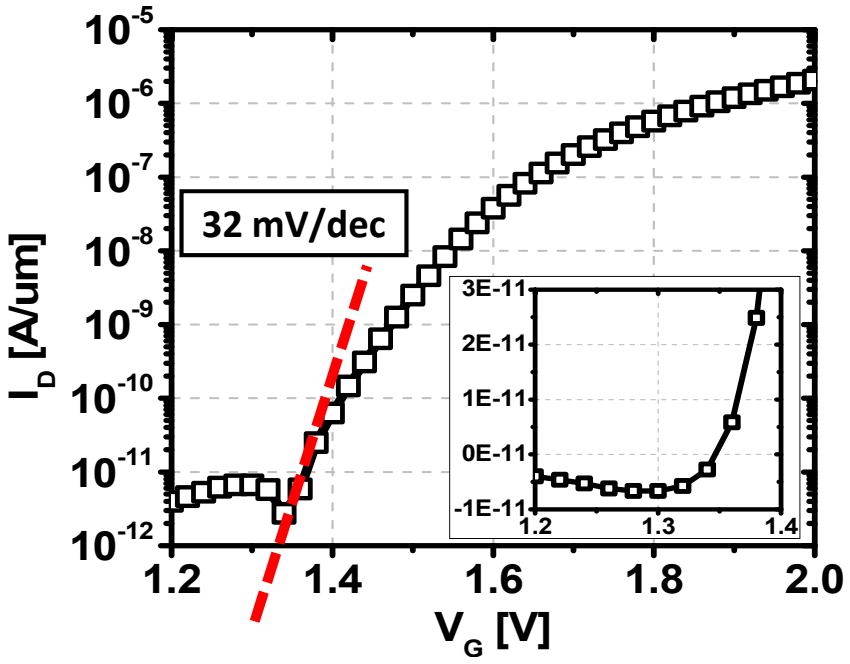


Figure 4.21 Zero crossing of drain current plotted in log scale can be mistaken as steep subthreshold swing

The subthreshold swing in this case can be mistaken to be 32mV/dec but when plotted in linear scale in is due to the zero crossing of the drain current due to gate leakage components in this case. Therefore, the current data near any zero crossing should be discarded when evaluating the subthreshold swing.

In the presence of gate leakage, the direction of the currents with respect to the bias polarity should be carefully examined and matched to ensure the measured current data is actual tunneling current. Figure 4.22 shows an illustration of two screening criteria for the drain current in the presence of gate leakage and a questionable case where the date should be discarded. In all circumstances the direction of the drain current should match the bias polarity of the drain to source bias. If this condition is not met, the measured drain current can be coming from the gate or the substrate and should not be evaluated as tunneling current happening at the source. In the presence of gate leakage the direction of the gate current should be in the same direction of the drain current to ensure that the gate current is not adding to the drain current. In the opposite case, the drain current can be a sum of the tunneling current and the gate current and can be measured to be larger than the actual value. A questionable case is also shown in Figure 4.22 where the gate current is flowing into the device and the drain current is flowing out. First of all, the direction of the drain current is not matching the drain to source bias and also since the drain current is in the opposite direction of the gate current, we cannot tell where exactly the drain current is coming from. This can happen in the presence of large gate leakage.

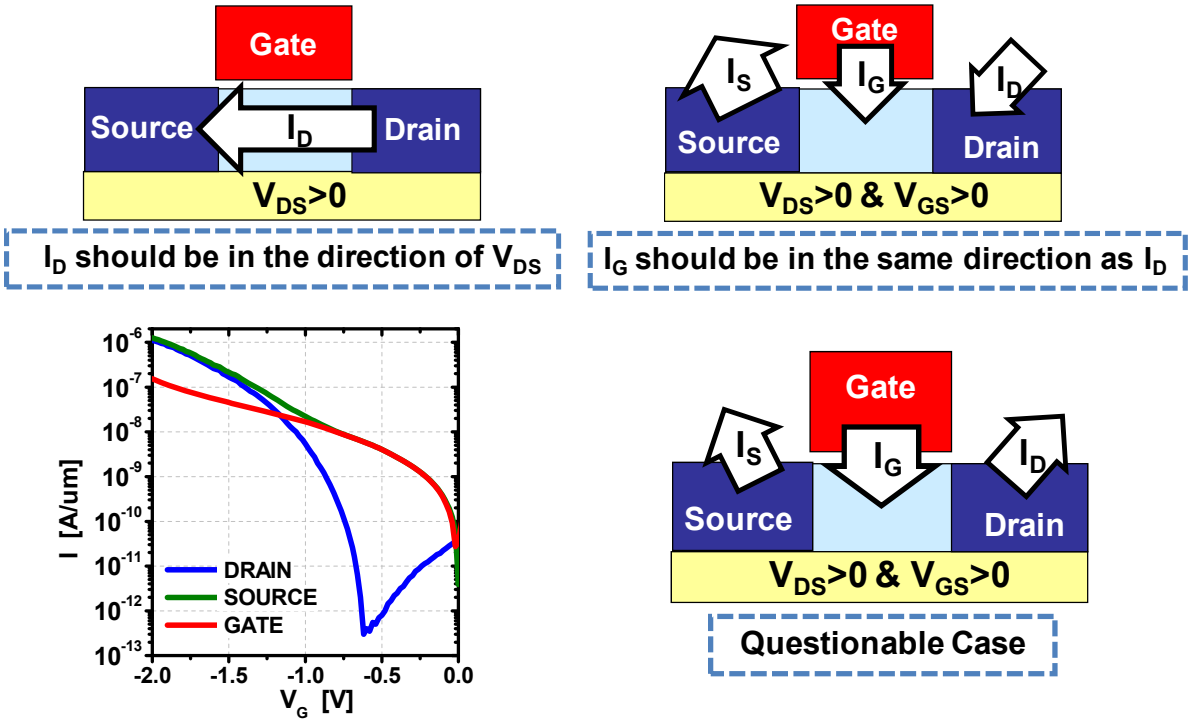


Figure 4.22 Data verification under gate leakage

In addition to the above criteria, the subthreshold swing, when plotted against the drain current, should be a smooth function. That is, it should be evaluated where the drain current is monotonically increasing or decreasing and not when it is fluctuating. Also, the subthreshold swing should not change with V_G sweep rate and direction, thus eliminating the possibility of the steep subthreshold swing coming from charging and discharging of traps or other potential wells in the device. Table 4.3 lists all the screening criteria that should be considered when evaluating the subthreshold swing in these steep subthreshold swing devices.

Value	Screen Criteria
I_D	Is I_D in the direction of V_{DS}?
	Is there a zero crossing?
I_G	Is it in the same direction as I_D?
SS	Is I_D monotonically increasing or decreasing?
	Is it consistent w.r.t. sweep time & direction?

Table 4.3 List of screening criteria for drain current, gate current and subthreshold swing (SS) to eliminate spurious data.

Figure 4.23 shows the statistical distribution of the minimum subthreshold swing of the silicided source TFET with and without the above screening applied. As can be seen in the figure, a much larger portion (>50%) of the devices show to have sub-60mV/dec subthreshold swing without the screening criteria applied where eliminating spurious data shows around 30%. Also, the subthreshold swing values can be mistaken to be steeper than they actually are.

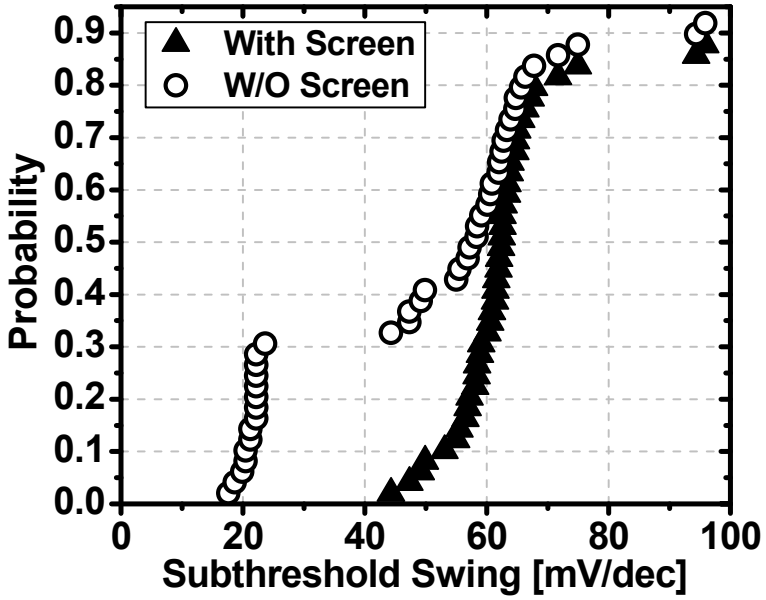


Figure 4.23 Subthreshold swing distribution with and without screening

4.5 Selective Silicide Using Germanium

In order for a more controlled process for the dopant segregation and the formation of the wedge shape required for the performance of this device, silicidation techniques using the difference in the silicidation rate of germanium compared to silicon were also explored. Selective silicide experiments were conducted on silicon wafers with blanket epitaxial layers of SiGe and Si on top. Following the epitaxial growth, the wafers were first etched down to the silicon substrate to form pillars of the Si/SiGe stack. Then nickel silicidation was performed to see the different silicidation rates of each layer. Figure 4.25 shows the cross section SEM image of the silicided Si/SiGe stack.

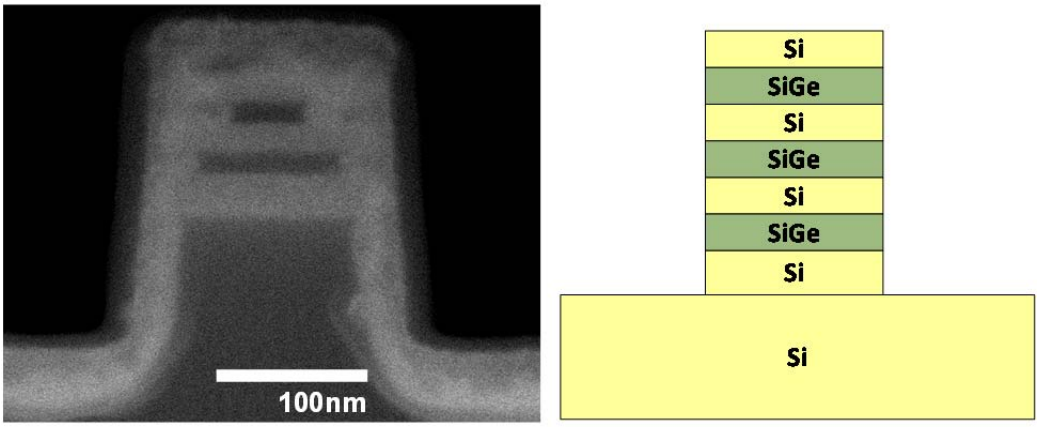


Figure 4.25 Cross-section SEM image of silicided Si/SiGe stack and structure of the Si/SiGe stack

As can be seen in the cross section SEM, the darker region which represent silicon shows more encroachment of the nickel silicide compared to the lighter germanium region. The upper layers shows to have faster silicidation rates, especially in the silicon layers and this is due to the fact that the upper layers have more defects or dislocations caused by the lattice mismatch of the SiGe to silicon. Utilizing this difference in silicidation rate of the two materials, the following structure and process flow will greatly improve the controllability of the silicide wedge pocket formation.

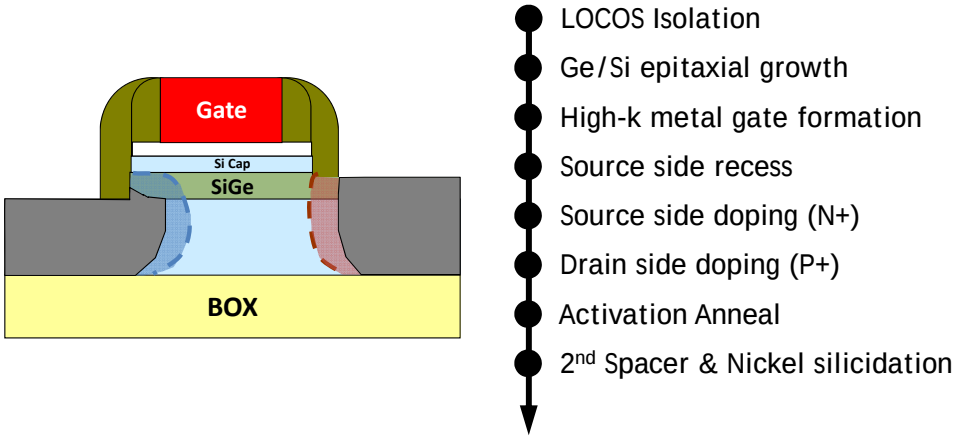


Figure 4.26 Proposed structure and process flow

The source and drain are formed asymmetrically to ensure the formation of the wedge only in the source region. After isolation, a thin (~3nm) epitaxial layer SiGe or Ge is grown on the active regions followed by a thin silicon cap (1~2nm) in order to improve the gate dielectric to channel interface quality. Then the gate stack and seal nitride is formed. The source side is recessed to expose the silicon substrate which would allow the silicide to form from the silicon substrate and grow upward to the channel surface where the thin SiGe or Ge layer will help slow down the silicidation rate making it more controllable. Then the source/drain are doped using ion implantation followed by thermal anneal. Different doping and anneal techniques will be required in order to limit the amount of Ge diffusion to the channel surface and ensure the gate dielectric to channel interface. Silicidation is done after a second spacer is formed.

4.6 Summary

A silicide source TFET exhibiting 46mV/dec subthreshold swing is demonstrated using dopant segregation with nickel silicide. A high I_{ON}/I_{OFF} ratio of 7×10^7 for 1V operation is measured. Statistical distribution of subthreshold swing show more than 30% of the device to have a sub-60mV/dec subthreshold swing whereas control devices without the dopant segregation process do not show any. Improvement in the subthreshold swing distribution is observed with high-k gate dielectric compared to SiO_2. Barrier height measurement of holes and electrons at the source confirm the existence of n-type dopant pile up at the silicide to silicon interface. Simulation shows that in addition to the abrupt profile from silicide induced dopant segregation, the wedge shape pocket of the silicide is crucial in the steep performance of the silicided source TFET. GIDL and GISL measurements show effectiveness of the source profile and wedge shape pocket. Nonlinear change in the subthreshold swing with respect to temperature is observed as expected. Simulation shows temperature dependent current characteristics change depending on the doping level at the region where tunneling is occurring due to the temperature dependence of the Fermi level.

Screening criteria in evaluating the current and subthreshold swing of TFETs are developed and applied to eliminate phantom data coming from leakage, zero crossing or transient effects. The subthreshold swing distribution without the screening applied gives a much higher percentage of the devices showing sub-60mV/dec subthreshold swing.

Selective nickel silicide using germanium is explored and is shown that the wedge shape pocket can be formed by utilizing the slower silicidation rate of SiGe. A silicided source TFET structure with this SiGe pocket and fabrication flow is proposed.

4.7 References

[4.1] M. Zhang, J. Knoch, Q.T. Zhao, U. Breuer, and S. Mantl, "Impact of dopant segregation on fully depleted Schottky-barrier SOI-MOSFETs," *Solid-State Electronics*, vol. 50, no. 4, Apr. 2006, pp. 594-600

[4.2] K. Jeon, W.-Y. Loh, P. Patel, C.Y. Kang. J. Oh, A. Bowonder, C. Park, C.S. Park, C. Smith, P. Majhi, H.-H. Tseng, R. Jammy, T.-J.K. Liu, and Chenming Hu, "Si tunnel transistors with a novel silicided source and 46mV/dec swing," *VLSI Technology, 2010 Symposium on*, pp. 121-122, June 2010

[4.3] W. Y. Choi, B.-K. Park, J. D. Lee, and T.-J. King Liu, "Tunneling Field-Effect Transistors (TFETs) With Subthreshold Swing (SS) Less Than 60 mV/dec", *IEEE Electron Device Letters*, vol. 28, no. 8, pp. 743-745, Aug. 2007

[4.4] F. Mayer, C. Le Royer, J.-F. Damlencourt, K. Romanjek, F. Andrieu, C. Tabone, B. Previtali, and S. Deleonibus, "Impact of SOI, $Si_{1-x}Ge_xOI$ and GeOI substrates on CMOS compatible Tunnel FET performance," *IEDM Technical Digest*, pp. 163-165, Dec. 2008

[4.5] T. Krishnamohan, D. Kim, S. Raghunathan, and K. Saraswat, "Double-Gate Strained-Ge Heterostructure Tunneling FET (TFET) With record high drive currents and <60mV/dec subthreshold slope," *IEDM Technical Digest*, pp. 947-949, Dec. 2008

[4.6] S. M. Sze and Kwok K. Ng, Physics of Semiconductor Devices, 3rd Edition, John Wiley & Sons Inc., New Jersey, 2007

[4.7] H. Iwai, T. Ohguro, S. Ohmi, "NiSi salicide technology for scaled CMOS," *Microelectronic Engineering*, vol. 60, no. 1–2, pp. 157-169, January 2002

[4.8] Taurus MEDICI User Manual, Version A-2007.12, Synopsys, Mountain View, CA

[4.9] C. Hu, D. Chou, P. Patel, A. Bowonder, "Green Transistor - A V_{DD} Scaling Path for Future Low Power ICs", *VLSI Technology, Systems and Applications (VLSI-TSA), 2008 International Symposium on*, pp. 14-15, April 2008

[4.10] Y. Yoon and S. Salahuddin, "Inverse Temperature Dependence of Subthreshold Slope in Graphene Nanoribbon Tunneling Transistors", *Applied Physics Letters*, vol. 96, pp. 013510, Jan. 2010

[4.11] Y. Yoon and S. Salahuddin, "Non-linear Temperature Dependence of Subthreshold Current in Tunneling FETs Based on Graphene Nanoribbon", *MRS Spring Meeting*, April 2010

Chapter 5

Improving the ON Current

5.1 Introduction

In the previous chapter, we have shown that using a dopant segregation technique with nickel silicide and a wedge shaped pocket formation with a source side recess we could achieve sub-60mV/dec subthreshold swing on silicon P-I-N TFETs. But the drive current measured on these devices are very low in the order of 10^{-6}A/μm. In order to compete with the conventional MOSFET, the drive current of these devices needs to be increased at least 2 orders of magnitude to be in the range of 100μA/μm. As shown in Chapter 2, the tunneling probability is exponentially dependent on the bandgap and since silicon is a relatively large bandgap material, the most intuitive method of increasing the tunneling current would be to reduce the bandgap. That is, to use a lower bandgap material. This approach will be discussed in Chapter 6. In this chapter, an attempt to increase the drive current of TFETs is discussed by way of utilizing a different structure and process techniques with silicon wafers. Two structures will be discussed. The first structure will be utilizing a larger gate to source overlap with high dopant activation using spike+flash anneal techniques [5.1~5.2] and the second structure utilizing an undoped vertical pocket concept.

5.2 P-I-N TFET with High I_{ON} using Spike+Flash Anneal

Since the band-to-band tunneling generation of electrons and holes at the source of a P-I-N TFET is similar to the GIDL (Gate Induced Drain Leakage) in a MOSFET [5.3], it would be desirable to design the P-I-N TFET with increased gate-to-source overlap in order to induce more tunneling. MEDICI simulations of P-I-N TFETs with varying gate-to-source overlap are conducted [5.4] and shown in Figure 5.1.

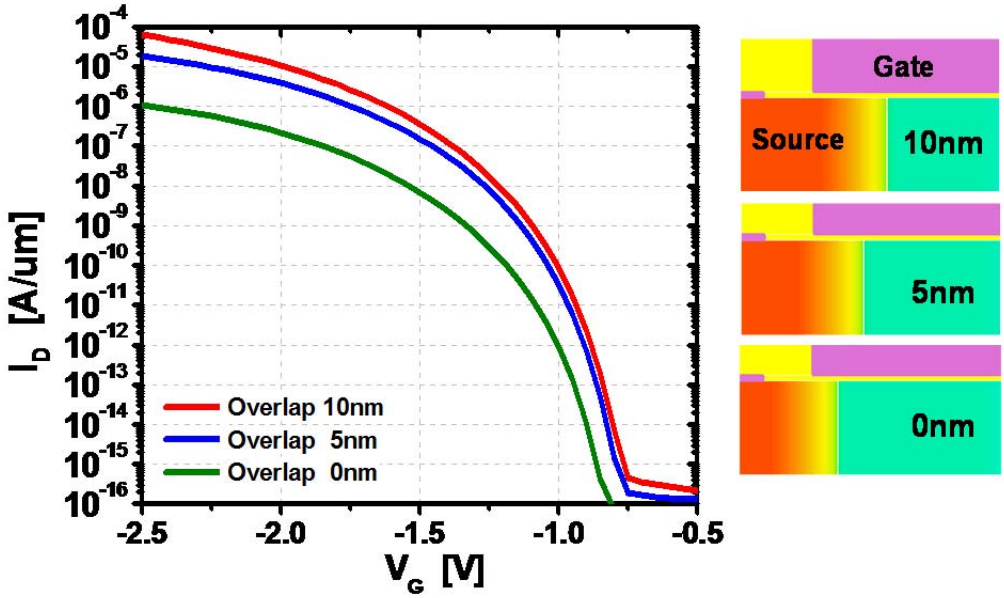


Figure 5.1 Simulated silicon P-I-N TFET I_D-V_G with various gate to source overlap. Simulation results show an increase in tunneling current with larger overlap where the overlap is defined as the length between the gate edge and source doping region not including the lateral gradient. (Source doping concentration $N_D=5\times10^{19}cm^{-3}$ with a doping gradient of 5nm/dec)

Simulation results show that the tunneling current increases with respect to an increase in the overlap between the gate and source due to the increase in the band-to-band tunneling generation area. The amount of increase is from 0nm overlap to 5nm overlap shows to be larger than that between 5nm and 10nm. This shows that the increase in the drive current will saturate above a certain amount of overlap. Although the band-to-band generation of carriers increases with the overlap, there is a limit in the efficiency that the carrier can be pulled out through the drain.

In order to increase the gate to source overlap, P-I-N devices were fabricated with intentional overlap by forming the source with the half mask before the gate stack. The simplest method to increase the overlap would be to subject the devices to higher thermal budget but this would degrade the source doping profile and only give us one split per wafer, not considering variation effects. Implanting and annealing the source with the half mask before gate stack formation allows us to have different gate to source overlap for different channel lengths. Figure 5.2 shows the process flow for the intentional overlap P-I-N TFET. After LOCOS isolation, the source is implanted (As, $2\times10^{15}cm^{-2}$, 10keV) using the half mask aligned to the center of the gate, giving an overlap of $L_G/2$ for each device. The source is then annealed using spike anneal after removing the implant mask. The high-k/metal gate stack is formed followed by seal nitride leaving the high-k foot to reduce gate dielectric edge leakage. Then the drain is implanted (B, $3\times10^{15}cm^{-2}$, 4keV) and annealed using millisecond flash anneal at 1200°C. Flash anneal was chosen to limit the amount of diffusion of the source dopants. Self-aligned silicide is formed after the second spacer. Then metallization was done up to M1 layer.

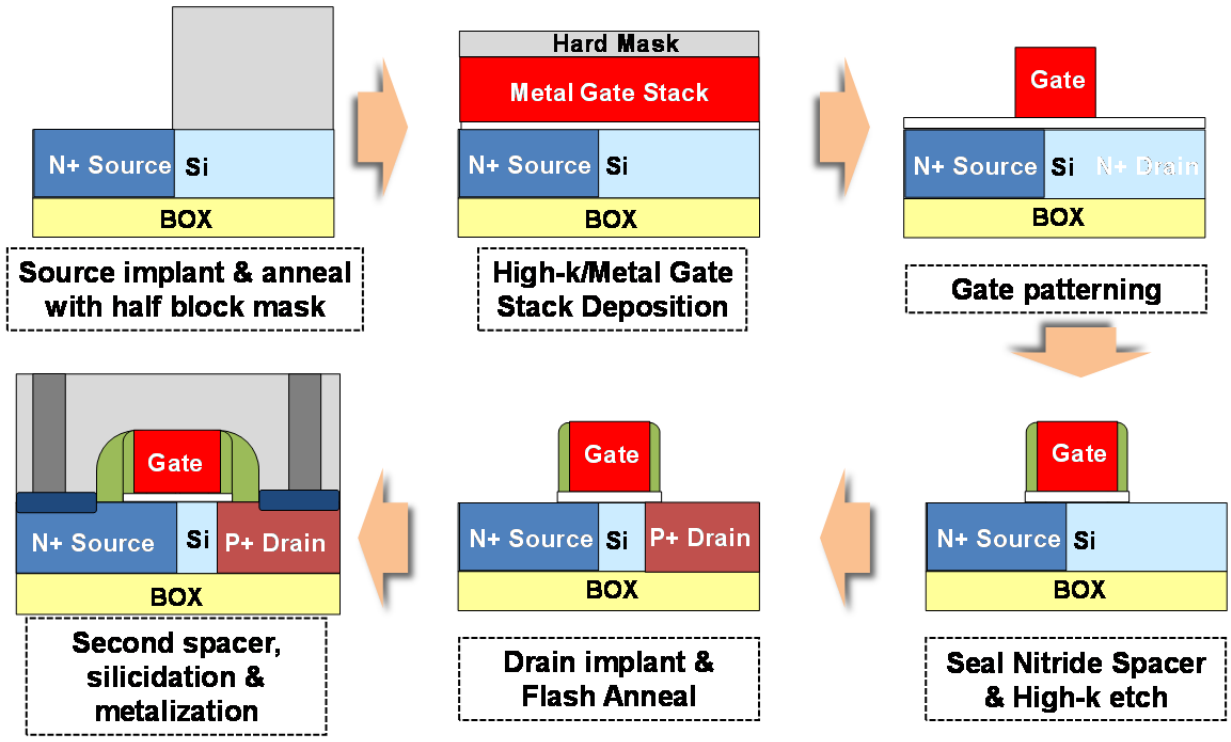


Figure 5.2 P-I-N TFET with intentional gate to source overlap process flow using spike+flash anneal combination

Figure 5.3 shows a high-resolution TEM image of a L_G=65nm device. The fabricated gate length is measured to be 56nm with a much shorter effective channel length due to the intentional overlap of the gate and source. The high-k foot can be observed to protrude beyond the gate edge.

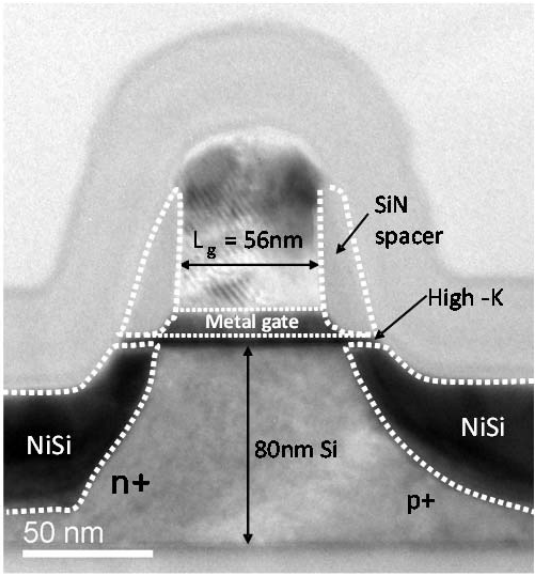


Figure 5.3 TEM image showing the cross-section of the fabricated device

Figure 5.4 shows the measured I_D-V_G and I_D-V_D characteristics of the 56nm gate length device. As can be seen in the I_D-V_G characteristics, a high drive current of >100μA/μm at an overdrive of 2.5V could be observed. The drive current is exceptionally high compared to previously fabricated and reported silicon TFETs [5.5~5.9] and this is due to the intentional overlap added by the effect of enhanced source dopant activation due to the spike+flash anneal combination.. This demonstrates that high ON current in TFETs can be achieved with optimal source doping and gate to source overlap. Gate leakage currents in these devices were measured to be more than three orders of magnitude lower than the drive current and hence not affecting the tunneling current. Ambipolar behavior in these devices is greatly suppressed owing to the asymmetry caused by the intentional gate to source overlap. NFET devices can also be fabricated using an intentional overlap of the P+ side. Although high drive currents could be achieved in these short channel devices, a sub-60mV/dec subthreshold swing could not be observed. The steepest subthreshold swing values measured in these devices were in the range of 100mV/dec. This is due to the fact that the source to drain leakage floor is high in these devices. As shown in the turn-on characteristic of the silicided source TFET in the previous chapter, silicon TFETs show the steep subthreshold swings at low current ranges and tend to degrade as the tunneling probability increases. So in order to observe the steep portion of the subthreshold swing, the leakage current needs to be suppressed. This is not easy to achieve in such short effective channel devices. This can also be observed in the non-saturating behavior of the I_D-V_D characteristics given in Figure 5.4 (b). The short effective channel length allows more control of the tunneling current by the drain. Another factor to the large subthreshold swing would be due to the less abrupt doping profile of the source formed using the conventional implant and anneal technique. This causes different doping concentration regions to turn on at different stages in the gate bias, resulting in a relatively gradual turn-on behavior.

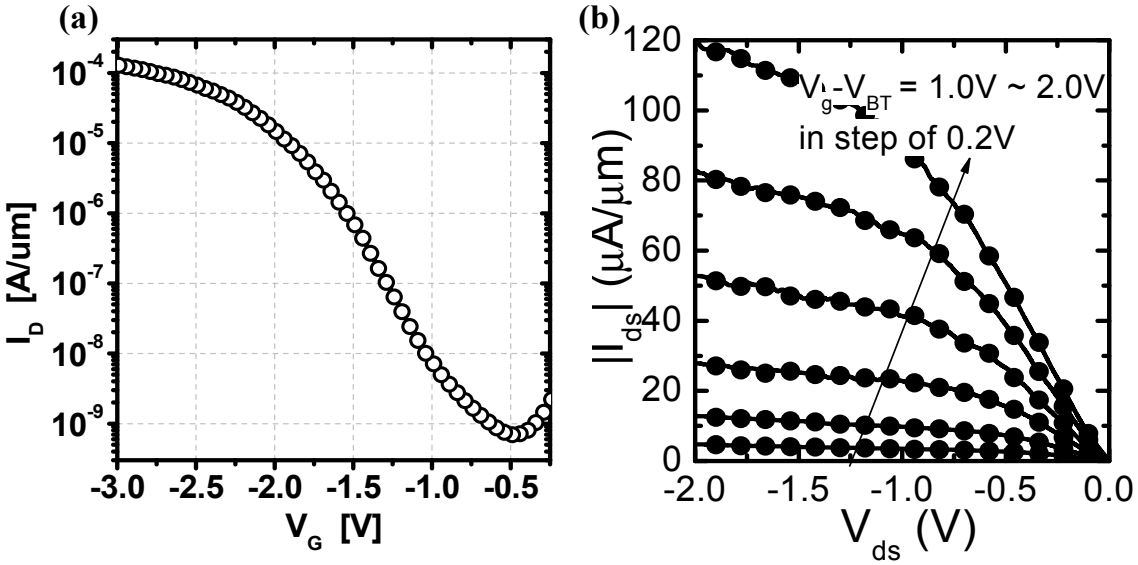


Figure 5.4 **(a)** Measured I_D-V_G characteristic of the intentional overlapped device with a gate length of 56nm. A high ON current of > 100μA/μm is measured at $|V_G-V_{BT}| = 2.5V$ and $V_{DS} = -1V$ **(b)** Measured I_D-V_D characteristics

Due to the non-self-aligned nature of the process and the short gate length of these devices, they would be more prone to be misaligned during the implant processes and form a MOSFET. To eliminate the possibility of a MOSFET, diode measurements were done. Figure 5.5 shows the P-I-N diode measurement between the source and drain with the gate open. The measurement shows a well behaved diode. This confirms that the N+ and P+ implant did not extend over to the drain and source respectively.

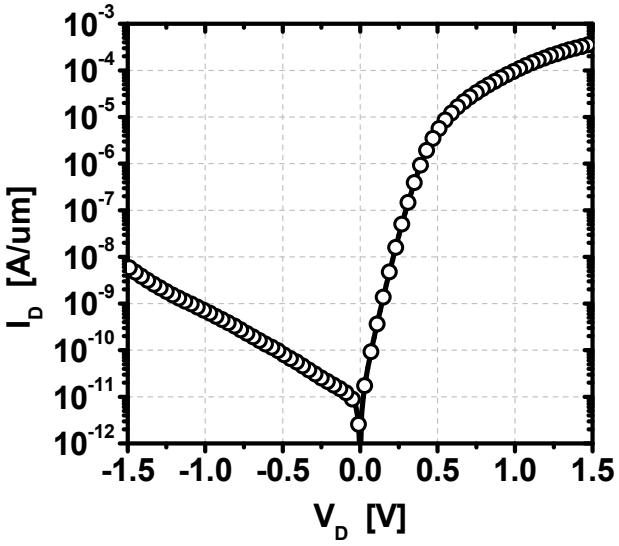


Figure 5.5 Diode measurement of the intentional overlapped P-I-N TFET showing a well behaved diode

Diode measurements with the gate biased at 0V were also conducted and an NDR (Negative Differential Resistance) behavior could be observed with a PVCR (Peak to Valley Current Ratio) of 1.36 at 213K. The NDR behavior observed in the diode confirms that band-to-band tunneling is occurring in the P-I-N structure.

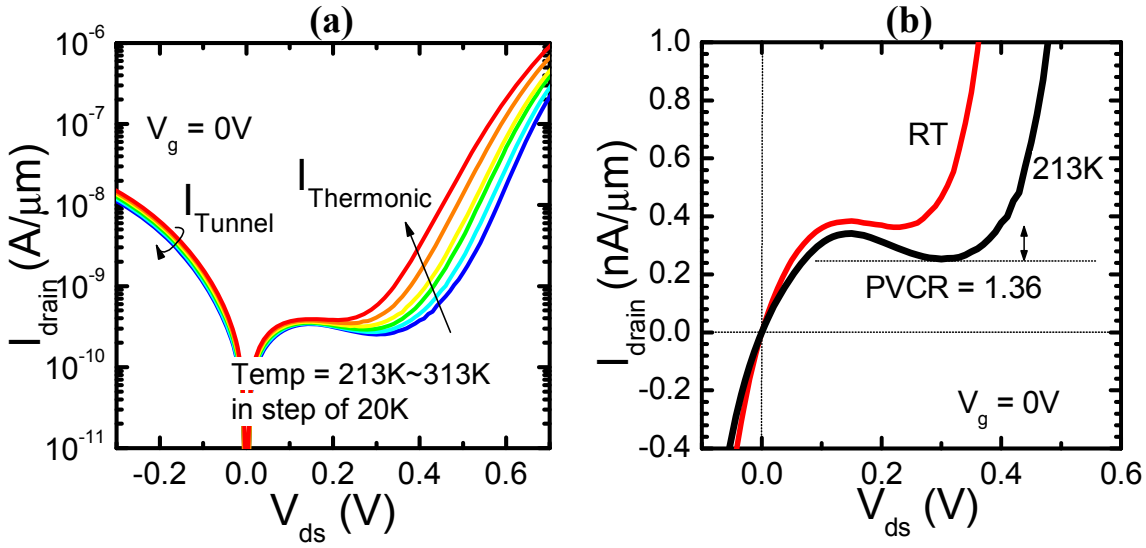


Figure 5.6 **(a)** Diode measurements at V_G=0V for temperatures 213K~313K show NDR behavior. **(b)** Room temperature and 213K compared in linear scale

Splits with spike anneal only were included to see the effect of the spike+flash anneal combination. The spike only devices were fabricated using a dummy gate process where the drain (P+) was implanted using the dummy gate. After the removal of the dummy gate, the source (N+) implant was done using the half mask followed by spike anneal to anneal both the source and drain. This split was not subject to the drain implant and flash anneal after the actual gate formation. Figure 5.7 compares the I_D-V_G and C-V measurement of the spike anneal only and spike+flash anneal splits. In the I_D-V_G comparison we can see that the spike+flash anneal sample shows a higher tunneling current. This shows that the additional flash anneal seen by the arsenic dopants in the source caused a higher activation rate thus allowing an increase in the tunneling probability. We can also observe that the benefit of the intentional overlap in the spike only sample by comparing with the P-I-N structure TFETs fabricated using conventional implant and anneal techniques given in chapter 3. The spike only sample with the intentional overlap shows a large improvement in the drive current. Higher leakage currents are measured in both the spike only and spike+flash sample due to the short effective channel length. Gate leakage currents in both devices are comparable and more than three orders of magnitude lower than the drain current. Both samples show very similar C-V behavior yielding almost the same EOT of ~1.1 nm and an interface trap density of $10^{11}cm^{-2}$. C-V measurements were taken at 100kHz with both source and drain grounded. This shows that the improvement in the drive current is not due to changes in gate control.

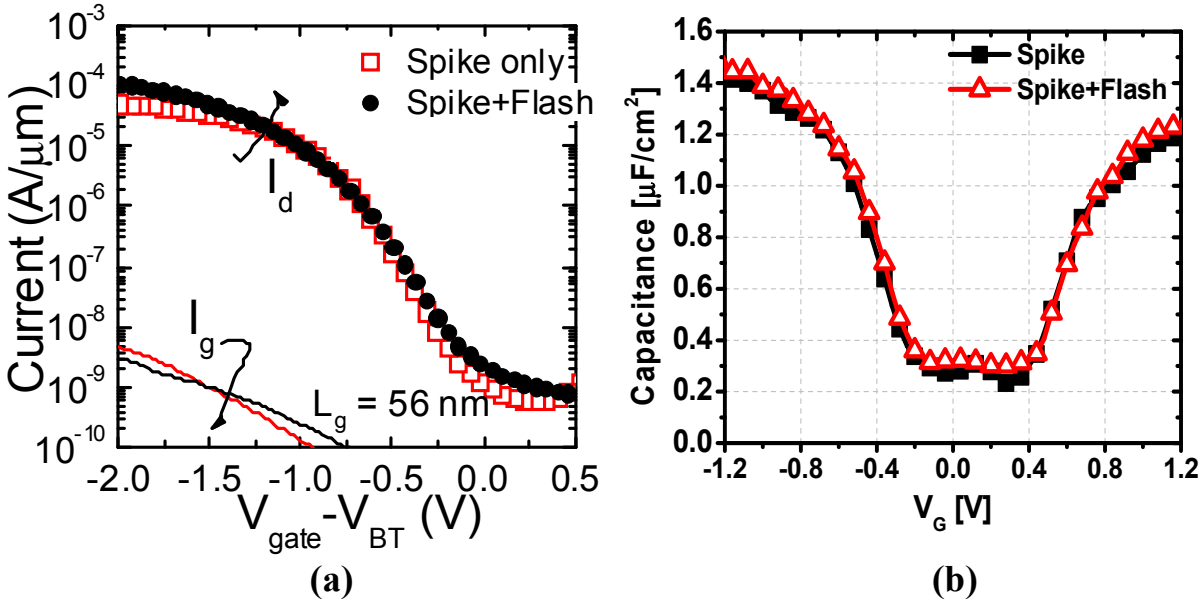


Figure 5.7 **(a)** I_D-V_G (V_{DS}=-1.0V) and **(b)** C-V (f=100kHz) measurement comparison of devices that were subject to spike only anneal and spike+flash anneal.

Statistical distribution of the subthreshold swing and I_{ON} vs. I_{OFF} are compared between the two samples and are given in Figure 5.8. The spike+flash anneal sample shows to have an overall lower subthreshold swing and better I_{ON} vs. I_{OFF} distribution. The large variation in the subthreshold swing and currents are expected to be due to the variation in the non-self-aligned process.

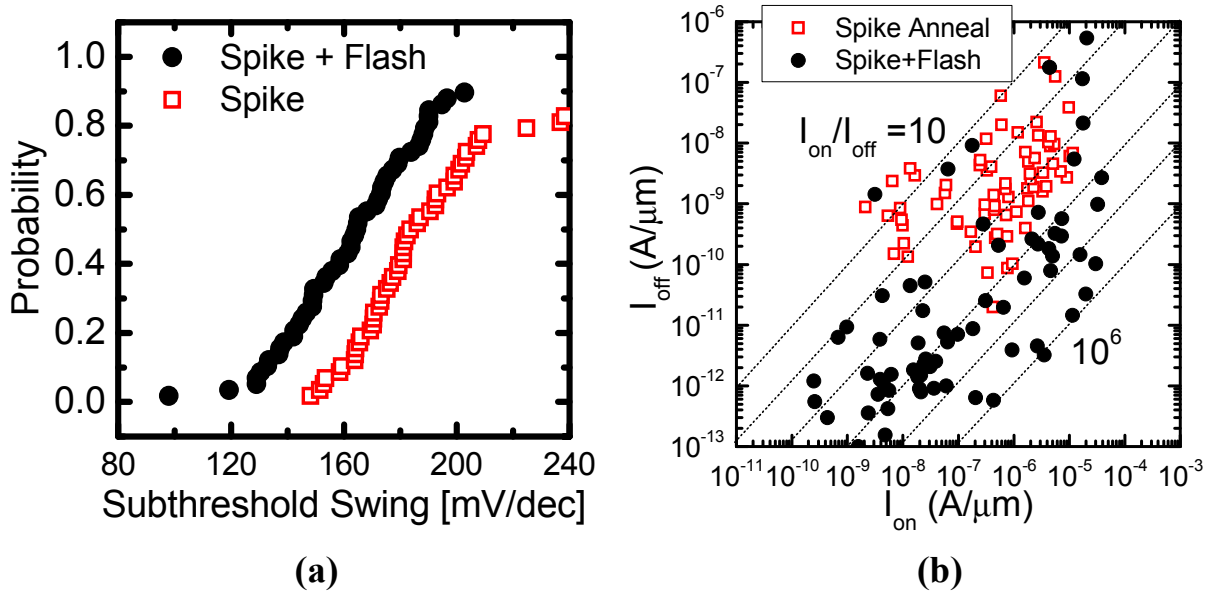


Figure 5.8 **(a)** Subthreshold swing distribution and **(b)** I_{ON} vs. I_{OFF} comparison of devices that were subject to spike only anneal and spike+flash anneal.

P-I-N TFET simulations were conducted with varying gate-to-source overlap and source doping concentration using the two-dimensional device simulator MEDICI. The results are shown in Figure 5.9 and compared with the measured results of the spike+flash anneal sample. Figure 5.9 (a) shows the results with varying gate-to-source overlap. Simulation results of a 10nm overlap agrees well with the simulation and suggests that the effective channel length of the measured device is approximately 46nm and the source implant half mask was misaligned approximately 18nm towards the source end. This agrees with the mean misalignment measured after the half mask exposure for the source which was 25nm towards the source. Figure 5.9 (b) shows simulation results with varying source doping concentration and overlap of 10nm. Simulation results with a concentration of $5\times10^{19}cm^{-3}$ agrees well with the measured data.

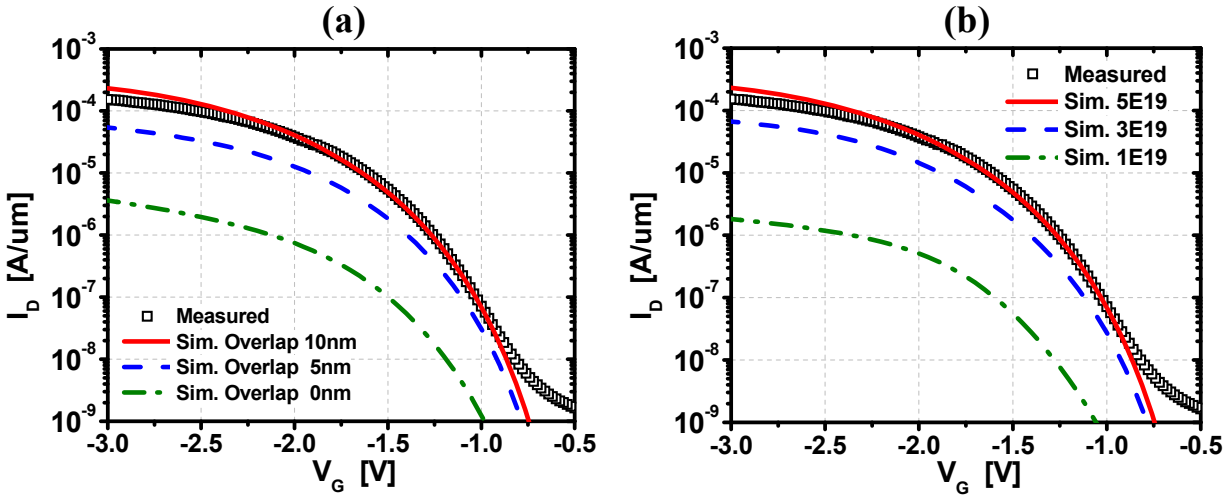


Figure 5.9 MEDICI simulation results of P-I-N TFET compared to measured data with varying **(a)** gate-to-source overlap (Source doping=5E19) and **(b)** source doping concentration (overlap=10nm)

Figure 5.10 show the temperature dependent characteristic of the spike+flash anneal sample. The ON current shows to be insensitive to temperature which confirms that the tunneling is occurring at a highly doped region where the Fermi level has a weak temperature dependence. On the other hand the leakage current increases exponentially with temperature due to thermal generation of carriers. The measured subthreshold swing shows to change in a nonlinear fashion to temperature as was observed in the silicide source TFET and predicted by atomistic simulation [5.10~5.11].

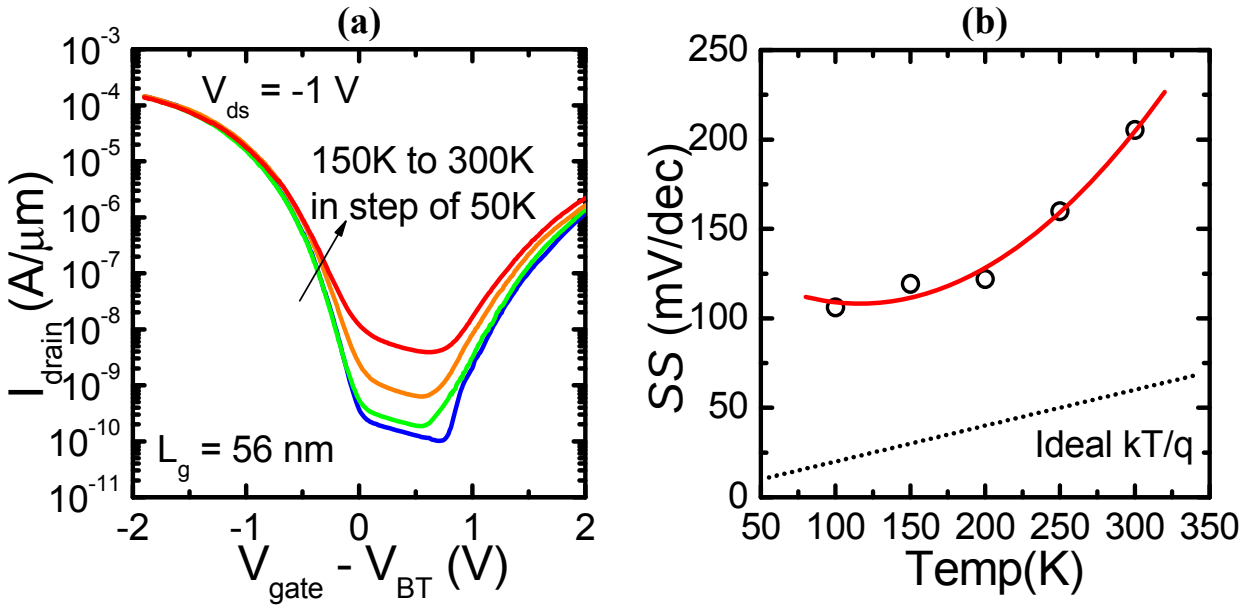


Figure 5.10 **(a)** Temperature dependent I_D-V_G measurements of the intentional overlapped P-I-N TFET with spike+flash anneal **(b)** Subthreshold swing vs. temperature showing a non-linear dependence

5.3 Undoped Vertical Pocket Structure TFET

The vertical pocket TFET or "Green" TFET (gTFET) adds an ultra thin, fully depleted highly doped pocket in addition to the intentional gate-to-source overlap and is shown to have superior turn-on characteristics compared to a general TFET [5.12~5.14]. The gTFET structure of an N-type device is shown in Figure 5.11 (a). This highly doped pocket allows for a uniform the band-to-band tunneling to occur vertically from the source to the pocket. Also the turn-on voltage of the gTFET can be adjusted with the dose of charge in the pocket just like a threshold voltage adjust implant with the tunneling current proportional to the pocket of overlap area. Figure 5.11 (b) shows the energy band diagram of the gTFET in the "ON" and "OFF" state. In the "OFF" state, band-to-band tunneling cannot occur since there is no overlap between the conduction band of the pocket and the valence band of the source. As a positive gate bias is applied, the gate

electric field pulls the bands down allowing for a sudden overlap to occur between the two bands and thus a sudden jump in the tunneling current. Simulated I_D-V_G of the gTFET is given in Figure 5.11 (c) with different pocket doping concentration. We can see that with increased pocket doping concentration the turn-on voltage shifts with improved turn-on characteristics.

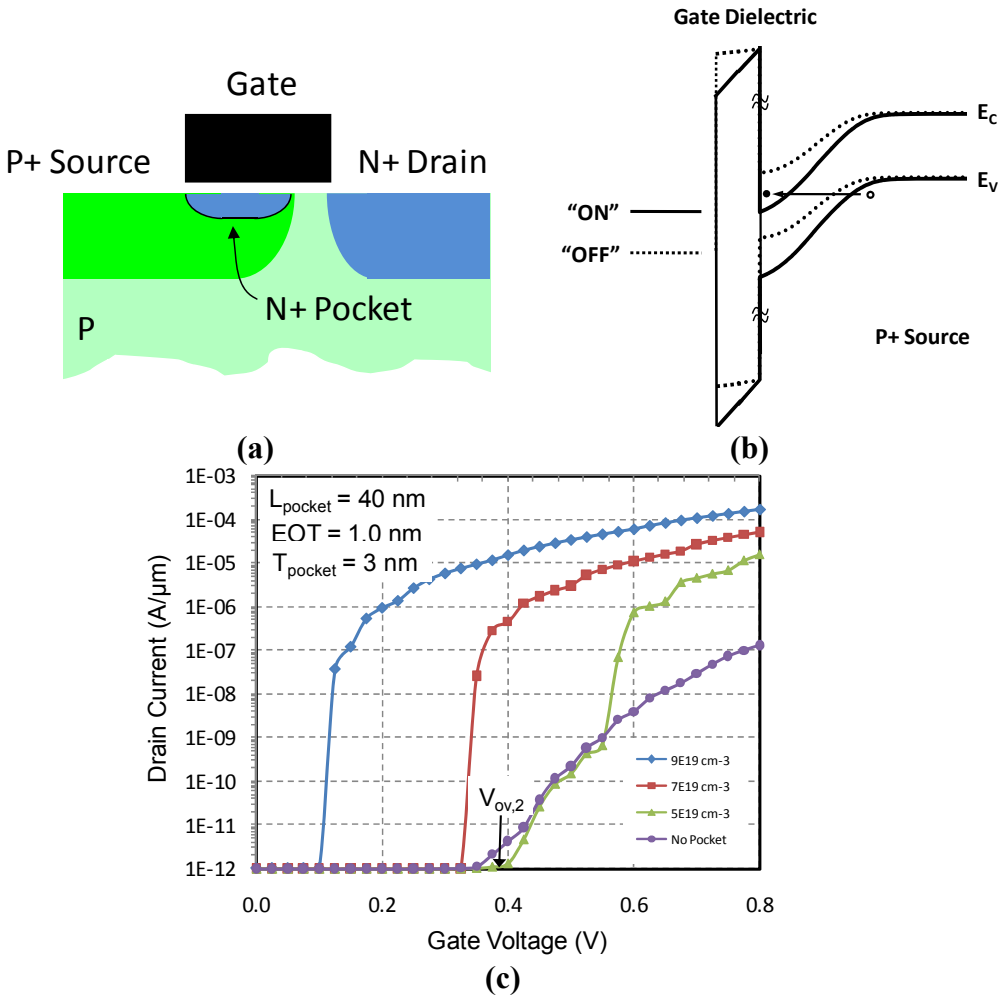


Figure 5.11 "Green" TFET **(a)** Structure of N-type device (b) Energy band diagram showing "ON" and "OFF" state (c) Simulated I_D-V_G [5.15]

To achieve these benefits of the gTFET, an ultra shallow, heavily doped pocket is required to be formed under the gate aligned to the edge of the source which is not easy to fabricate. In order for a more simpler process, an undoped vertical pocket structure is proposed and is shown in Figure 5.12 (a). The undoped pocket can be epitaxially grown on top of the source using a gate last process. MEDICI simulation results (Figure 5.12 (b)) show that improved turn-on characteristics can be achieved compared to the non-pocket device (0nm) and optimal thickness of the pocket is 7nm.

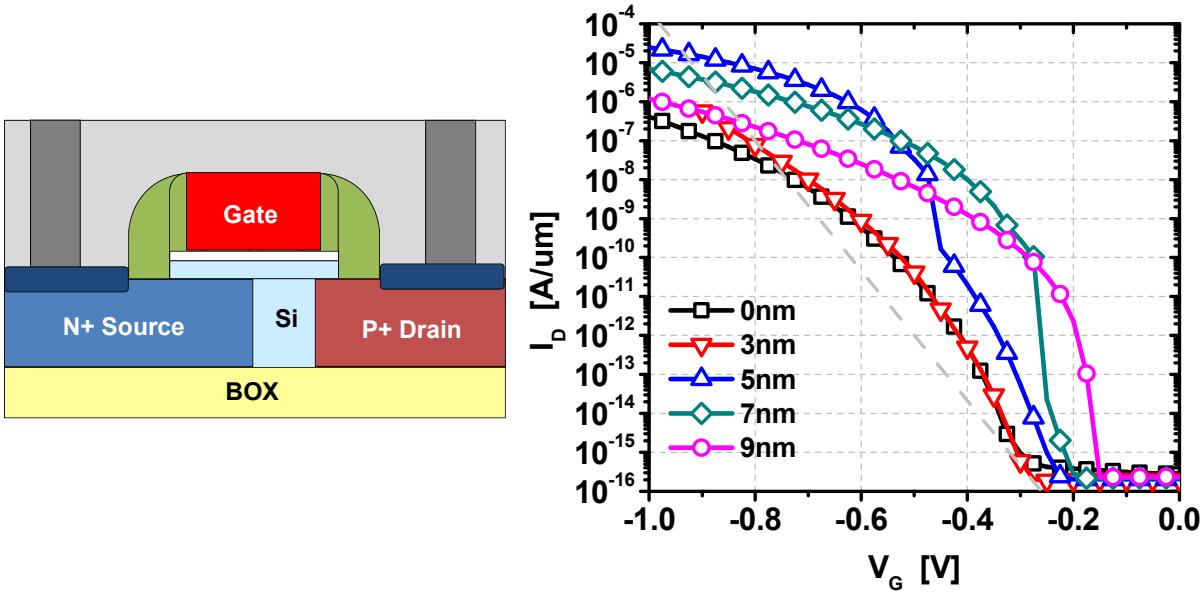


Figure 5.12 Undoped vertical pocket TFET **(a)** Structure **(b)** Simulated I_D-V_G

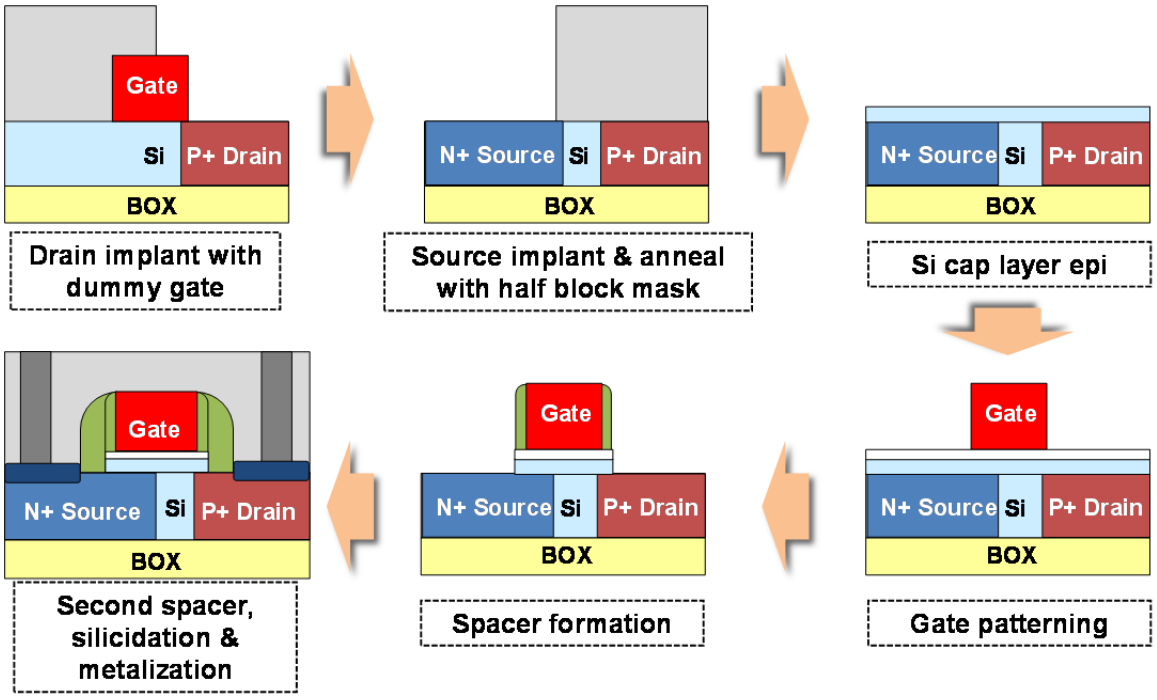


Figure 5.13 Process flow of the undoped vertical pocket TFET

Figure 5.13 shows the fabrication process flow of the undoped vertical pocket TFET. After LOCOS isolation, a nitride dummy gate is formed and the P+ drain is implanted (B, 3×10^{15}cm^{-2}, 4keV) aligned to the edge of the dummy gate using the half implant mask. Then the dummy gate

is etched away followed by source implant using the reverse half implant mask aligned to the center of the gate, giving an overlap of $L_G/2$ for each device. Both source and drain are annealed at the same time using spike anneal after removing the implant mask. Then a thin silicon cap was epitaxially grown on top of the active region to form the undoped pocket. 5nm and 7nm cap thickness were included as splits and a control without the silicon cap was also included. After the undoped pocket growth, high-k/metal gate stack is formed followed by seal nitride leaving the high-k foot to reduce gate dielectric edge leakage. During the seal nitride etch, the silicon cap on the source and drain regions are also etched away and a second nitride spacer was formed in order to remove the direct connection of the silicide to the pocket. Self-aligned silicide was formed with nickel using a two-step process and then metallization was done up to M1 layer.

For the undoped silicon pocket epitaxial growth, growth recipes were developed for crystalline silicon growth on silicon substrate in a CVD tool. Since the grown silicon cap cannot be distinguished from the substrate, patterns were formed on deposited oxide followed by epitaxial growth of silicon. Growth time and temperature was tuned to meet the thickness target. Figure 5.14 shows results of the epitaxial growth experiment on bulk wafers. As can be seen from the figure, 7nm, 10nm and 12nm of silicon cap was successfully grown with reasonable thickness uniformity. The surface roughness also shows to be smooth. The tuned recipes for 5nm and 7nm silicon cap growth were used in the actual device fabrication.

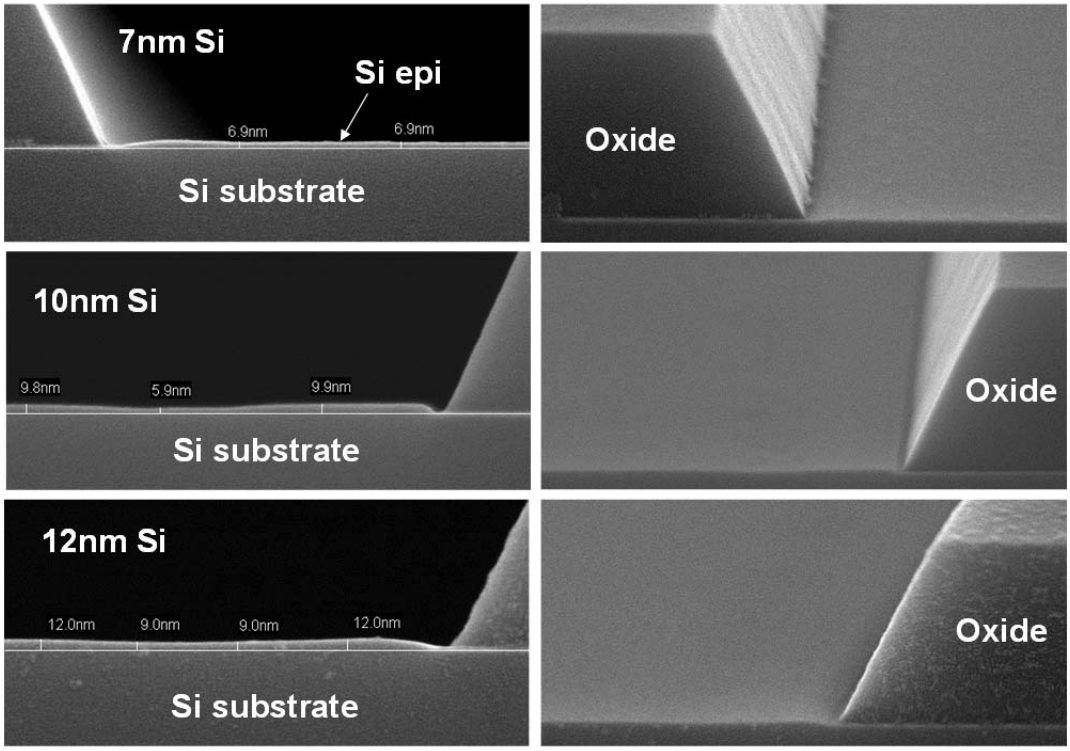


Figure 5.14 Cross-section SEM of epitaxial silicon growth on silicon substrate showing growth results for 7nm, 10nm and 12nm cap thickness.

Figure 5.15 shows the measured I_D-V_G and CV curves of the undoped vertical pocket TFET devices. Devices with 0nm, 5nm and 7nm pocket thickness are compared. Improvement in the drive current that was expected from simulation results could not be observed in the I_D-V_G measurement of 65nm drawn gate length devices. All three devices show similar current characteristics. Variation in the OFF state current or the turn on of tunneling in the P+ drain can be seen but do not show a trend against the pocket thickness. This can be expected to be due to the variation in the alignment of the gate to drain from the gate last process and variation in the P+ doping profile of the drain. Differences in the capacitance versus gate voltage characteristics could not be observed as well. This suggests that the epitaxial silicon cap was not grown properly or diffusion of the source dopants occurred rendering the role of the undoped pocket to be invisible. The same EOT of approximately 1.1nm was measured for all devices.

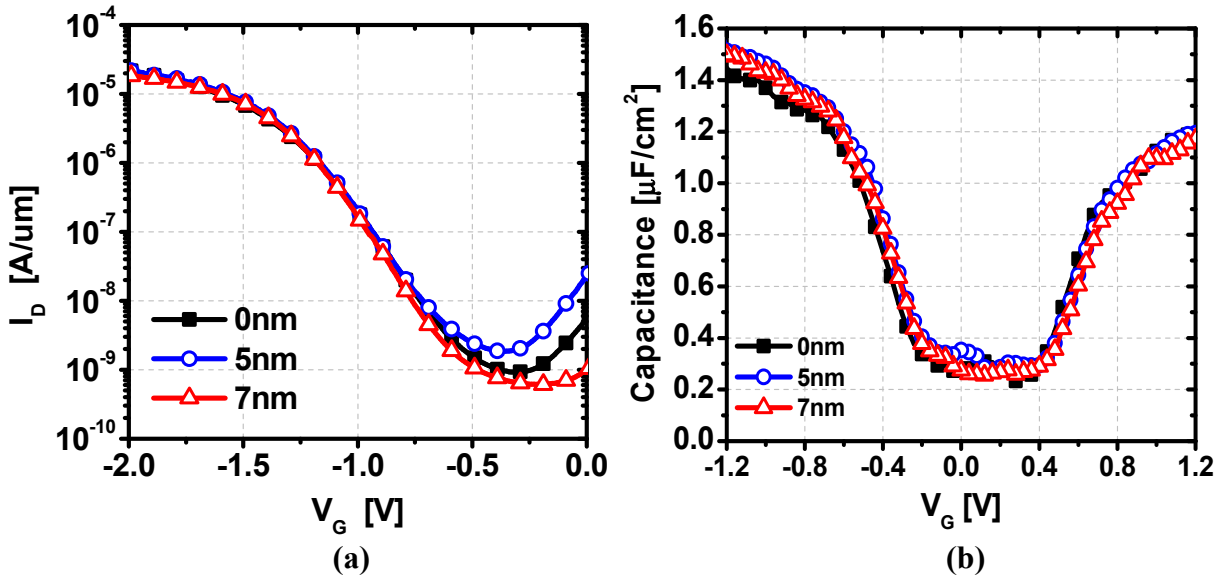


Figure 5.15 **(a)** Measured I_D-V_G of undoped vertical pocket TFET (L_G=65nm) **(b)** C-V measurement comparing 0nm, 5nm and 7nm pocket thickness.

Improvement in the ON current distribution could not be observed as well. Figure 5.16 (a) shows the ON current distribution of 65nm gate length devices for each wafer with 0nm, 5nm and 7nm pocket thickness. All three wafers show similar range (shaded box indicates 25%~75% percentile) of ON currents. The mean values of the ON current for all three wafers were between 2~5μA/μm with the 0nm control wafer showing the highest. Similar subthreshold swing distribution could be seen in the three splits as will with no devices having a subthreshold swing below 60mV/dec. The lowest subthreshold swing measured was approximately 80mV/dec on a 5nm thick pocket device. Other than the lowest few points the similar distribution in the subthreshold swing could be observed. (Figure 5.16 (b))

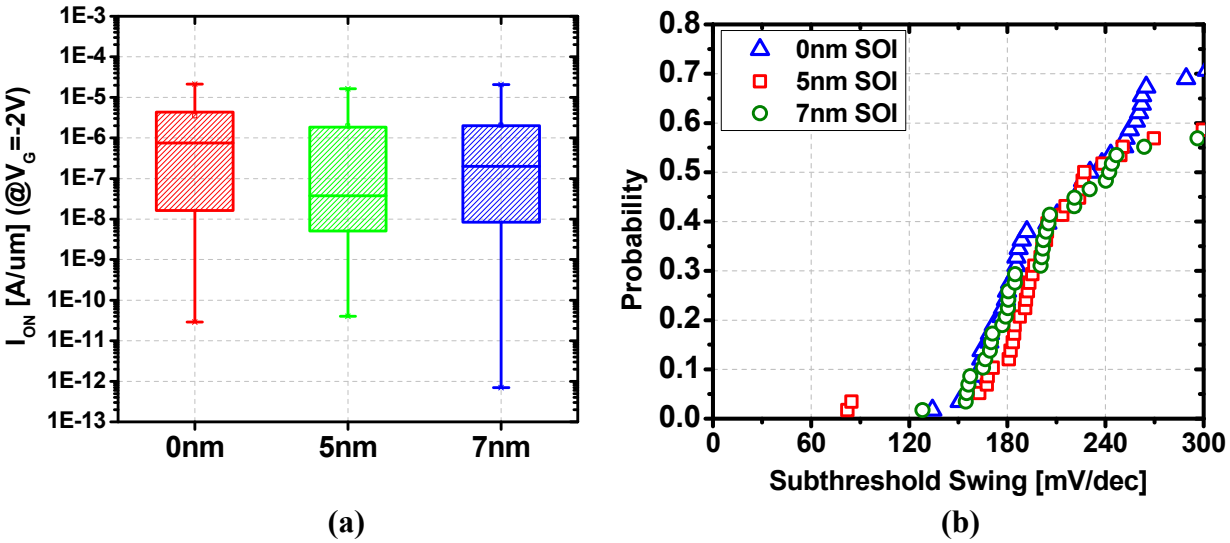


Figure 5.16 **(a)** ON current and **(b)** subthreshold swing distribution of undoped vertical pocket TFET with 0nm, 5nm and 7nm pocket thicknesses

In order to analyze the devices further, cross-section SEM images were taken. Below shows the image for the device with 7nm thick pocket device. It turns out that we could not measure any increase in the substrate thickness in the channel region below the gate. The SOI wafers initially had 80nm of silicon on oxide but the final device with the 7nm silicon epitaxial growth recipe used, shows about 80nm of substrate thickness. The source and drain regions are much thinner due to the over-etch done after the seal nitride spacer to remove the grown silicon cap in these regions. We speculate the growth conditions developed for the oxide openings in the test wafers were directly not applicable to growth in the active regions due to the difference in size. Silicon growth experiments on actual device wafers should be done and analyzed to improve the process.

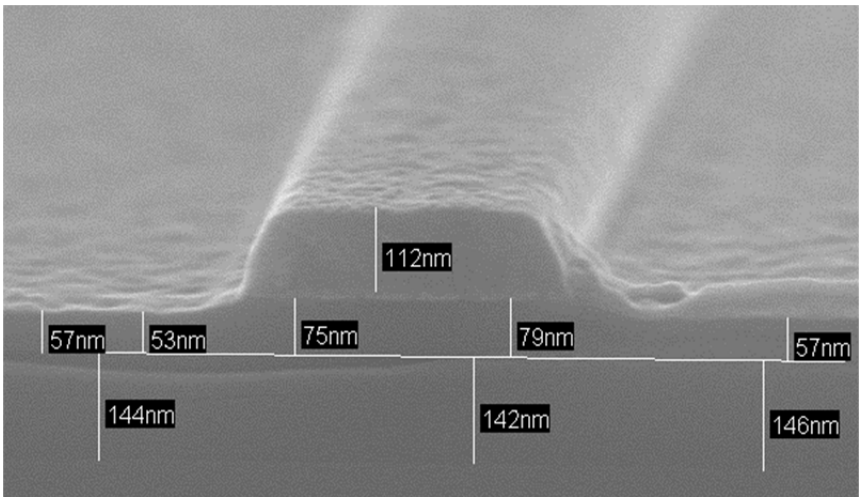


Figure 5.17 Cross-section SEM of undoped vertical pocket TFET. Thickness increase due to the silicon cap growth cannot be seen.

5.4 Summary

P-I-N TFETs with intentional gate to source overlap were fabricated by using a gate last process and spike+flash anneal combination. High tunneling currents of >100μA/μm at an overdrive of 2.5V were demonstrated on 56nm gate length devices. Diode measurements were done to confirm the alignment of the half mask of the short channel devices. Rectifying behavior is observed with gate open and an NDR behavior is measured with V_G=0V confirming band-to-band tunneling. Comparison to devices with only spike anneal shows higher dopant activation is achieved and thus higher drive current due to the spike+flash anneal combination. Improved subthreshold swing and I_{ON}/I_{OFF} distribution is observed. MEDICI simulation results with 10nm gate-to-source overlap and a source doping concentration of $5\times10^{19}cm^{-3}$ show good agreement with measured data. The subthreshold swing of the intentional overlapped P-I-N TFET shows to have a nonlinear temperature dependence.

Undoped vertical pocket structure TFETs were fabricated as an alternative to the vertical pocket gTFET. The undoped pocket allows for a more controllable process with the control of the pocket thickness done through epitaxial growth of silicon. Simulation results show 7nm is optimal thickness of undoped pocket. A gate last process was used for the fabrication of the undoped vertical pocket TFET with 0nm, 5nm and 7nm thickness splits. Electrical measurements show no difference in the current or capacitance characteristics between the thickness splits. Physical analysis confirms the undoped silicon epitaxial growth process did not deposit any silicon. More development in the silicon growth is required.

5.5 References

[5.1] W.-Y. Loh, K. Jeon, C. Y. Kang, J. Oh, P. Patel, C. Smith, J. Barnett, C. Park, T.-J. King Liu, H.-H. Tseng, P. Majhi, R. Jammy and C. Hu, "Sub-60nm Si Tunnel Field Effect Transistors with I_{ON} > 100μA/μm", *Proceedings of ESSDERC*, pp. 162-165, Sept. 2010

[5.2] W.-Y. Loh, K. Jeon, C. Y. Kang, J. Oh, T.-J. King Liu, H.-H. Tseng, W. Xiong, P. Majhi, R. Jammy, and C. Hu, "Highly scaled (L_g ~ 56nm) gate-last Si tunnel field-effect transistors with I_{ON} > 100μA/μm", *Solid-State Electronics*, vol. 65–66, pp. 22-27, Nov.–Dec. 2011,

[5.3] T.Y. Chan, J. Chen, P. Ko, and C. Hu, "The impact of gate-induced drain leakage current on MOSFET scaling", *IEDM Technical Digest*, vol. 33, pp. 718-721, 1987

[5.4] Taurus MEDICI User Manual, Version A-2007.12, Synopsys, Mountain View, CA

[5.5] W. Y. Choi, B.-K. Park, J. D. Lee, and T.-J. King Liu, "Tunneling Field-Effect Transistors (TFETs) With Subthreshold Swing (SS) Less Than 60 mV/dec", *IEEE Electron Device Letters*, vol. 28, no. 8, pp. 743-745, Aug. 2007

[5.6] F. Mayer, C. Le Royer, J.-F. Damlencourt, K. Romanjek, F. Andrieu, C. Tabone, B. Previtali, and S. Deleonibus, "Impact of SOI, $Si_{1-x}Ge_xOI$ and GeOI substrates on CMOS compatible Tunnel FET performance," *IEDM Technical Digest*, pp. 163-165, Dec. 2008

[5.7] T. Krishnamohan, D. Kim, S. Raghunathan, and K. Saraswat, "Double-Gate Strained-Ge Heterostructure Tunneling FET (TFET) With record high drive currents and <60mV/dec subthreshold slope," *IEDM Technical Digest*, pp. 947-949, Dec. 2008

[5.8] S. H. Kim, H. Kam, C. Hu and T.-J. King-Liu, "Germanium-Source Tunnel Field Effect transistors with Record High I_{ON}/I_{OFF}", *VLSI Technology, 2009 Symposium on*, pp. 178-179, June 2009

[5.9] K. Jeon, W.-Y. Loh, P. Patel, C.Y. Kang. J. Oh, A. Bowonder, C. Park, C.S. Park, C. Smith, P. Majhi, H.-H. Tseng, R. Jammy, T.-J.K. Liu, and Chenming Hu, "Si tunnel transistors with a novel silicided source and 46mV/dec swing," *VLSI Technology, 2010 Symposium on*, pp. 121-122, June 2010

[5.10] Y. Yoon and S. Salahuddin, "Inverse Temperature Dependence of Subthreshold Slope in Graphene Nanoribbon Tunneling Transistors", *Applied Physics Letters*, vol. 96, pp. 013510, Jan. 2010

[5.11] Y. Yoon and S. Salahuddin, "Non-linear Temperature Dependence of Subthreshold Current in Tunneling FETs Based on Graphene Nanoribbon", *MRS Spring Meeting*, April 2010

[5.12] C. Hu, "Green transistor as a solution to the IC power crisis," *9th International Conference on Solid-State and Integrated-Circuit Technology ICSICT*, pp. 16-20, 2008.

[5.13] A. Bowonder, P. Patel, K. Jeon, J. Oh, P. Majhi, H. H. Tseng, and C. Hu, "Low-voltage green transistor using ultra shallow junction and hetero-tunneling," *International Workshop on Junction Technology*, pp. 93-96, May 2008.

[5.14] P. Patel, K. Jeon, A. Bowonder, and C. Hu, "A Low Voltage Steep Turn-Off Tunnel Transistor Design," *International Conference on Simulation of Semiconductor Processes and Devices*, pp. 1-4, Sept. 2009.

[5.15] P. Patel, "Steep Turn On/Off "Green" Tunnel Transistors," Ph.D. thesis, Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, pp. 42-45, 2010

Chapter 6

Si/Ge Hetero-structure TFETs

6.1 Introduction

It was shown in Chapter 2 that the tunneling generation rate of germanium is many orders higher than silicon due to its lower bandgap. Processing of germanium has challenges such as achieving a good quality gate dielectric interface, dopant activation and ohmic contact to name a few. But germanium has been extensively researched as a replacement MOSFET channel material due to its higher mobility and many solutions to these processing challenges have been reported [6.1~6.4]. Due to these benefits, utilizing germanium is explored to improve the performance of TFETs. Pure germanium devices as well as strained and unstrained silicon/germanium hetero-structure TFETs are discussed.

6.2 Germanium P-I-N Structure TFET

In order to see the benefits in the tunneling current by using germanium, simple germanium P-I-N structure TFETs were fabricated. First a relaxed epitaxial layer of germanium (30nm) was grown on SOI wafers after isolation. A thin silicon cap of 1~2nm is grown on top of the germanium layer for improved gate dielectric interface quality. Then the high-k/metal gate stacks are deposited and patterned followed by seal nitride spacer formation. After the source and drain ion-implantation was done using the half mask aligned to the center of the gate, dopant activation was done at a lower 600°C for 60 seconds to limit the amount of germanium diffusion. For a simpler process and quicker turn around germanide and metallization steps were not done. Figure 6.1 shows the structure and cross-section SEM image of the fabricated germanium P-I-N TFET.

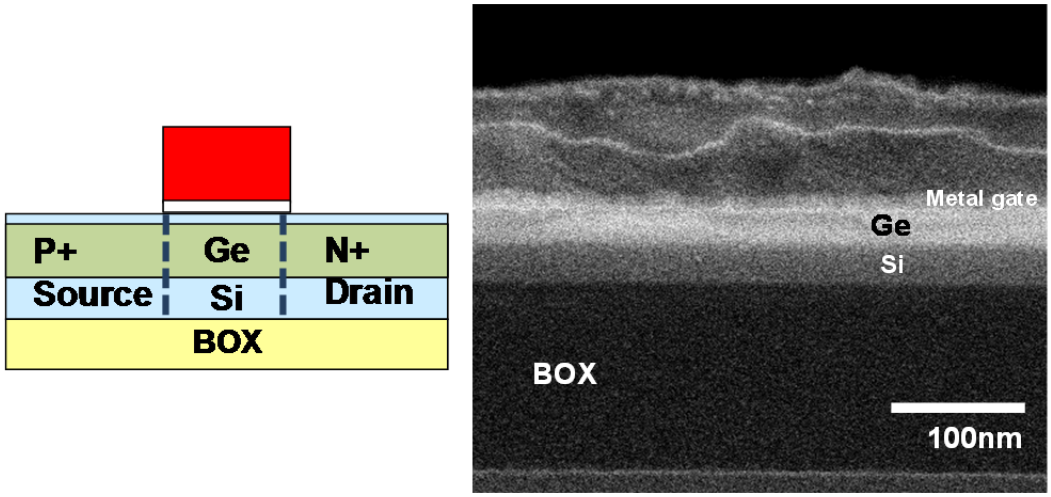


Figure 6.1 **(a)** Germanium P-I-N TFET structure **(b)** Cross-section SEM image of germanium P-I-N TFET showing epitaxial germanium layer between the high-k/metal gate stack and silicon substrate

Figure 6.2 (a) shows the measured I_D-V_G of the germanium P-I-N TFET. Silicon baseline P-I-N measurements are also given for comparison. Compared to the silicon device, an order of magnitude improvement in the drive current can be observed for an overdrive of 1.0V (=V_{GS}-V_{BTBT}). But a much larger increase in the leakage current is also observed. An increase in the junction leakage is expected due to the reduced bandgap and defects/dislocations caused by the relaxation of the germanium layer. Due to the larger leakage current, worse subthreshold swing and ON/OFF ratio is observed.

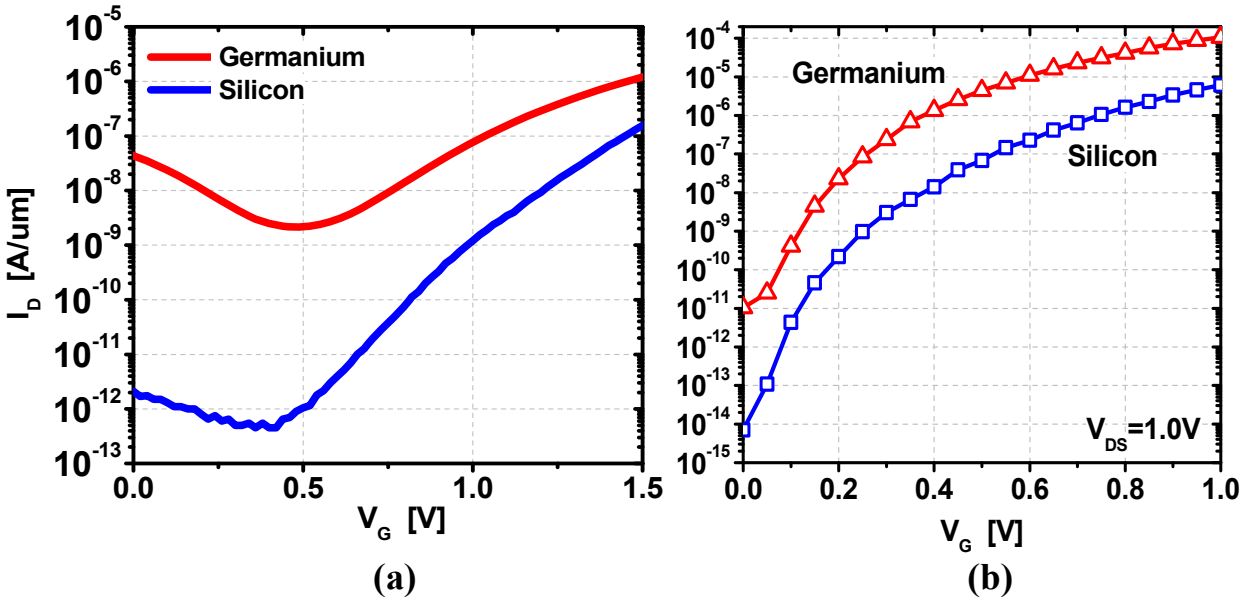


Figure 6.2 **(a)** Measured I_D-V_G of germanium P-I-N TFET compared to silicon **(b)** MEDICI simulation results comparing silicon and germanium P-I-N TFETs

MEDICI simulation of P-I-N TFETs using modified model parameters from Chapter 2 show similar trends (Figure 6.2 (b)). The drive current shows an order improvement but a much larger (three orders) increase can be seen in the OFF state. The degradation in subthreshold swing is also observed in the simulation results. Deviation in actual values compared to simulation is thought to be due to differences in doping profiles used in simulation and those achieved from ion implantation and anneal process. As in the case of silicon P-I-N TFETs, conventional ion implantation and anneal process for germanium P-I-N TFET would not be optimal and thus gives less drive current than is expected from simulation. Structures to improve the source doping by using in-situ doping during the growth of germanium and even lower tunneling bandgap by utilizing strain is discussed in the following sections.

6.3 Si/Ge Hetero-structure TFET

Since the tunneling occurs mainly in the source to channel interface in the P-I-N structure TFET, reducing the bandgap at this region would be most important for an increase in drive current. Thus we explore ways of using silicon and germanium hetero-structure at the source to channel interface and keep more silicon in the structure to utilize the developed baseline processes for silicon. Section 6.3.1 explores a structure where the source region in a silicon P-I-N is replaced with germanium which gives the benefit of the lower germanium bandgap in the tunneling process but also keeps the benefit of good gate dielectric interface of silicon. This would be the N-type device and Section 6.3.2 explores the complementary P-type device with the same silicon-germanium hetero-structure by utilizing germanium in the channel and silicon at the source.

6.3.1 Germanium Source Hetero-structure TFET

Figure 6.3 shows the structure of the germanium source hetero-structure TFET and the energy band diagram of the hetero-structure at the interface.

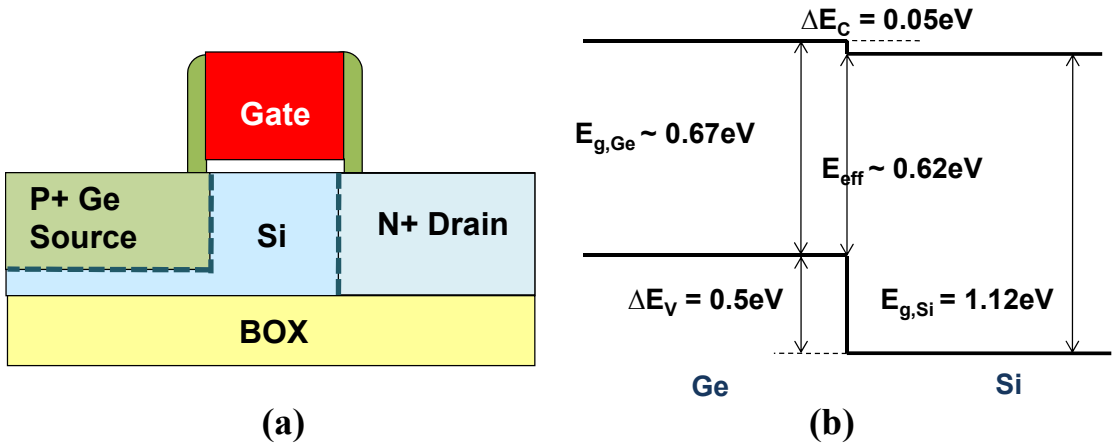


Figure 6.3 **(a)** Structure and **(b)** Energy band diagram of germanium source hetero-structure TFET

The germanium in the source is expected to be regrown after removing the silicon in the source region and doped to P+ as well during the growth process by using in-situ doping of boron. This would give an almost ideal step junction profile of the P+ drain which also aligns with the hetero-junction. The latter is more critical in achieving the increase in tunneling probability. If the dopant junction was not aligned with the hetero-junction and be in the silicon channel, which would be more likely to occur than the junction being formed inside the germanium source, all the tunneling happens in the silicon channel and we would not get the benefit of increased tunneling probability of the lower bandgap germanium. The band alignment of silicon and germanium gives a 50meV reduction in the effective tunneling bandgap which is measured from the valence band of germanium (source) to the conduction band of silicon (channel) due to the difference in the electron affinity of the two materials.

As in the case of silicon P-I-N TFETs, some overlap between the gate and source is helpful in increasing the tunneling current. So an isotropic etch process needs to be developed that can remove the silicon in the source as well as under the gate. Etch recipes were developed to achieve a rounded shaped trench with gate overlap by using a two-step etch process. Since more vertical etch depth is needed compared to the lateral etch distance, a vertical anisotropic etch is first done followed by an isotropic etch. Different etch techniques using in-situ etching by using a high dose arsenic implant and anneal at the source and etching during the plasma pre-clean step in the epitaxial growth CVD chamber were also developed and explored. Figure 6.4 (a) shows the etched profile of the source using this etch recipe. A rounded shape etch profile with gate undercut could be achieved. The drain side is protected with an oxide hard mask but half of the poly-silicon gate is exposed during the etch causing some removal of the gate. Germanium deposition will occur in the exposed gate region as well. Selective growth only in the source and exposed gate region is required to prohibit the source from shorting to the gate. Figure 6.4 (b) shows the germanium regrown in the etched trench on a test wafer where both the source and drain region were etched for a quicker turn around. The trench regions are fully filled with the germanium epitaxial layer without any voids.

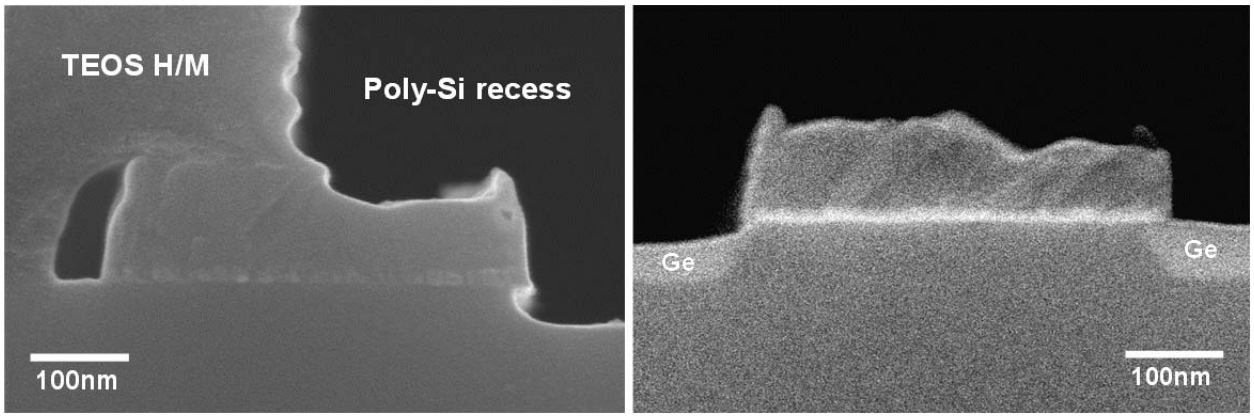


Figure 6.4 Cross-section SEM image of **(a)** isotropic source etch development and **(b)** Germanium epitaxial re-growth process development

Using the developed source etch and germanium growth recipes, germanium source hetero-structure P-I-N TFETs were fabricated. The baseline process developed for the silicon P-I-N was used to form the gate stack and drain. The interface oxide (SiO_x) was expected to see some damage during the isotropic etch, so a thicker than usual high-k gate dielectric was used for concerns of gate leakage. Then an oxide hard mask aligned to the center of the gate is used to protect the drain during the source etch and regrowth process with in-situ boron doping of the source. 25% and 50% germanium content SiGe was also included as splits to compare the benefit of the reduced effective tunneling bandgap as well as a silicon control P-I-N. Silicide step was not done due to process complexity coming from having different materials in the source and drain. Metallization up to M1 layer was done without the silicide and measured.

Although a thicker than usual high-k gate dielectric (HfO_2) was used, high gate to source leakage was measured in all devices. The gate leakage was so high that all currents underwent sign changes (zero crossing) and the tunneling current could not be extracted from the source/drain and gate terminal currents measured. Figure 6.5 shows a TEM image of the germanium hetero-structure TFET at the source region with the damaged gate dielectric region highlighted. The high-k gate dielectric showing up as the thin black line is missing at the gate edge. More etch process development is needed to test the etch rate of HfO_2 during source etch and may require a thicker or even a different material as the gate dielectric to eliminate this damage. The darker source region shows germanium growth and has some defects due to the relaxation of germanium.

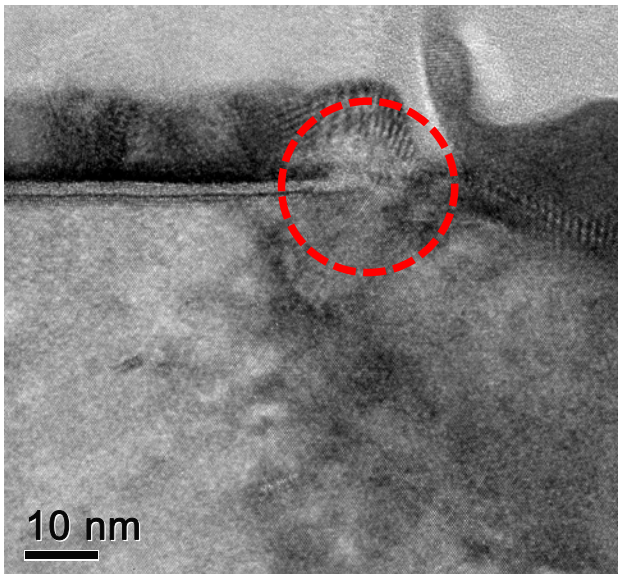


Figure 6.5 TEM image of germanium source hetero-structure TFET showing damaged gate dielectric at the edge region

6.3.2 Germanium Channel Hetero-structure TFET

A complementary P-type device of the germanium source hetero-structure TFET was also explored and fabricated. Figure 6.6 shows the structure and energy band diagram at the source of the germanium channel hetero-structure TFET. The effective tunneling bandgap is the same as the N-type device but gate control may be reduced due to germanium being used in the channel. So a thin silicon cap in the channel is needed to improve the gate dielectric interface quality. This brings the benefit of reducing the amount of gate dielectric damage during the source etch if a selective etch process is used to remove the germanium in the channel for gate to source overlap.

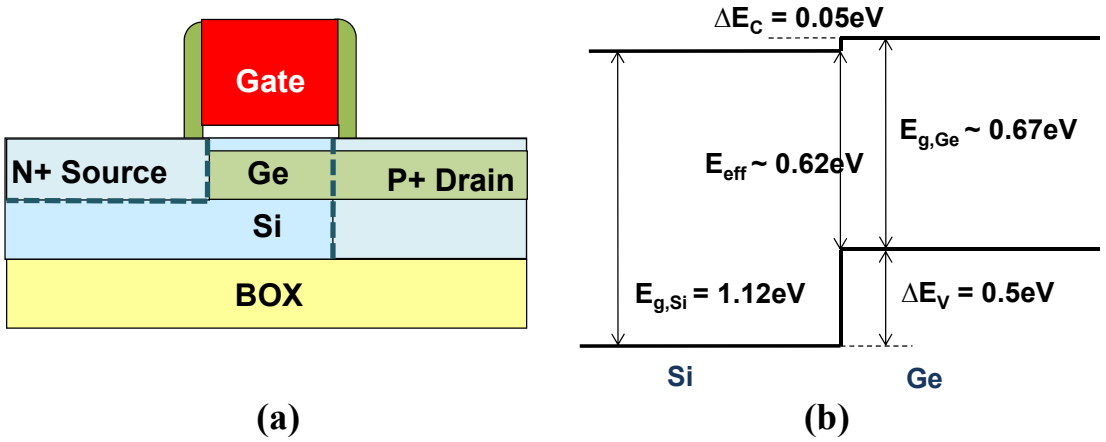


Figure 6.6 **(a)** Structure and **(b)** Energy band diagram of germanium channel hetero-structure TFET

Etch recipes were developed for selective etching of germanium sandwiched between two silicon layers. Figure 6.7 shows the selective etch results of 40nm Ge and 10nm $Si_{0.2}Ge_{0.8}$ with 5nm silicon cap. It can be seen that Ge and SiGe can be successfully etched away while leaving the silicon cap and layer underneath intact.

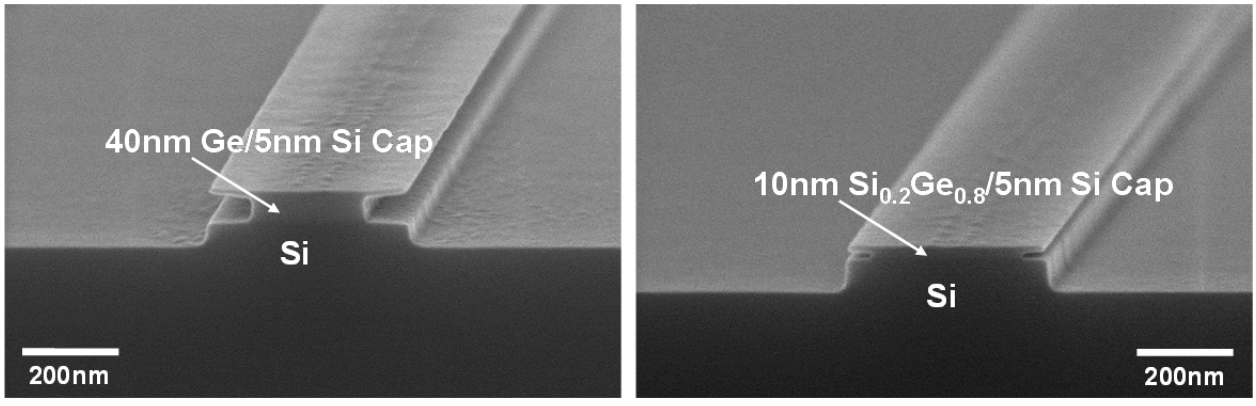


Figure 6.7 Selective etch process development. Selective etch of **(a)** 40nm germanium layer and **(b)** 10nm $Si_{0.2}Ge_{0.8}$ sandwiched between silicon

The germanium channel hetero-structure TFET process flow is based on the previously fabricated germanium P-I-N TFET. Relaxed germanium of 30nm is grown isolated active regions. Then the gate stack is formed followed by selective etching of the silicon cap and germanium is in the source region. Germanium is undercut to allow overlap of the regrown silicon source. Then the source is implanted with arsenic (As, 10keV, $5\times10^{15}cm^{-2}$) and drain with boron (BF_2, 5keV, $3\times10^{15}cm^{-2}$). In-situ doping of the N+ source could not be in this process. N-type in-situ doping during the epitaxial growth of silicon using phosphine caused unwanted etch removal of the grown silicon layer inhibiting the growth of silicon. The source and drain are annealed at 600°C for 60sec to limit the diffusion of germanium. This anneal condition would limit the amount of dopant activation in the source but using the high anneal temperatures for silicon would cause melting of the germanium. The wafers were pulled out and measured without silicide or metallization so that further anneal would be possible after preliminary measurement analysis.

Unfortunately, similar high gate to source leakage behavior of the germanium source TFET could be observed. Cross-section image analysis shows that the selectivity in the regrowth of silicon in the source was very poor allowing deposition of silicon on the side walls of the silicon nitride as well. This caused the source to form a short with the gate. Figure 6.8 shows the cross-section SEM image showing the bridge. We were not able to analyze the tunneling current in this device as well. More process development is needed for selectivity improvement of silicon growth on silicon compared to silicon nitride.

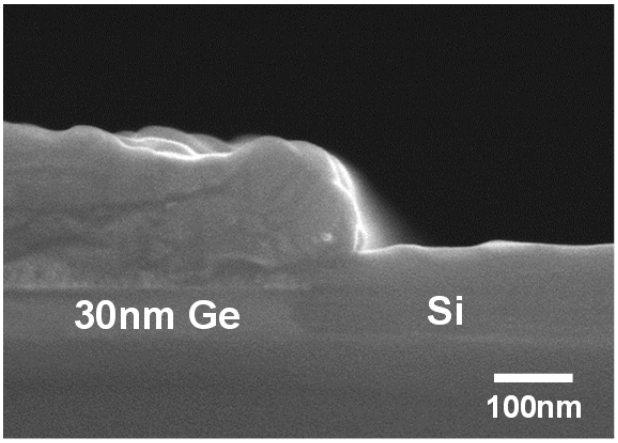


Figure 6.8 Cross-section SEM image showing silicon growth on nitride spacer sidewalls as well as source region

6.4 Si/Ge Strained Hetero-structure TFET

When silicon is grown lattice matched to germanium, a downward shift in the energy bands of the conduction band occurs reducing the bandgap of silicon to be smaller than the bandgap of germanium. In addition, the band alignment of strained silicon to relaxed germanium becomes favorable such that the effective bandgap measured from the valence band of germanium to the conduction band of strained silicon becomes even lower. Figure 6.9 (a) shows the band alignment of strained silicon on germanium. The reported bandgap of strained silicon on germanium and the effective tunneling bandgap (E_{Eff}) is given in Table 6.1. Reported values show that E_{EFF} can be as low as 0.16eV. This is lower than half the bandgap of InAs (E_g=0.36eV) and one fourth the bandgap of germanium. A vertical tunneling device utilizing this low effective tunneling bandgap can be fabricated (Figure 6.9 (b)) and will be able to achieve high tunneling currents.

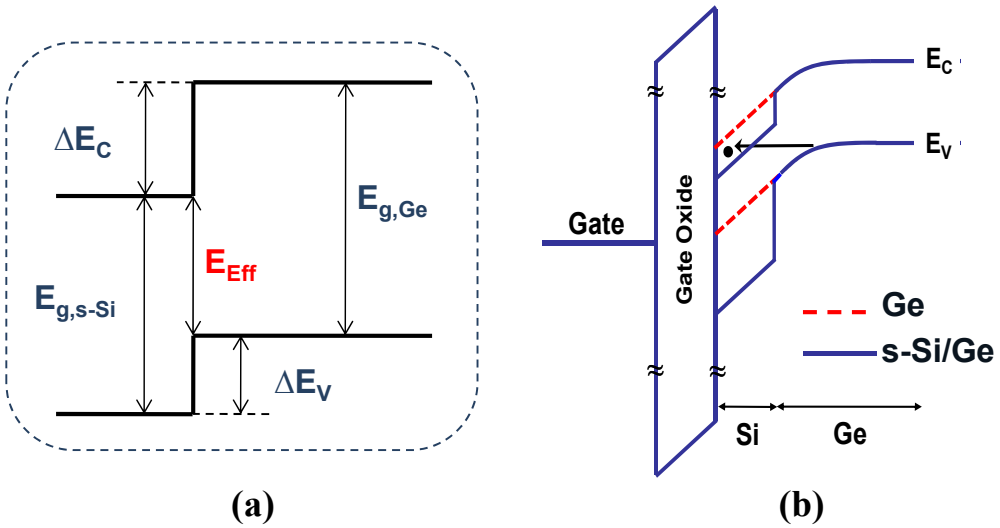


Figure 6.9 **(a)** Energy band diagram of biaxially strained silicon grown lattice matched to relaxed germanium **(b)** Energy band diagram of vertical tunneling structure using strained silicon (s-Si) on germanium

[001]	E_{Eff}	$E_{g,s\text{-}Si}$	ΔE_C	ΔE_V
Van de Walle et al [6.5]	**0.19**	0.42	0.55	0.31
Rieger et al [6.6]	**0.17**	0.35	0.57	0.25
Yang et al [6.7]	**0.16**	0.37	0.58	0.21
Schwartz et al [6.8]	n/a	n/a	n/a	0.17±0.13
Colombo et al [6.9]	n/a	n/a	n/a	0.21
Band Structure Lab [6.10]	**0.36**	0.54	0.30	0.18

[Units: eV]

Table 6.1 Band alignment of strained silicon on germanium

In the case of biaxially strained germanium grown lattice matched to silicon, an upward shift in the bands occur allowing for a small E_{Eff} as low as ~0.3eV to be achieved. Figure 6.10 shows the band alignment with reported s-Ge bandgap and E_{Eff} listed in Table 6.2. Orientation dependent band alignment calculation of s-Si on Ge and s-Ge on Si using Band Structure Lab shows [001] surface orientation to give lowest E_{EFF} (Table 6.3, Table 6.4).

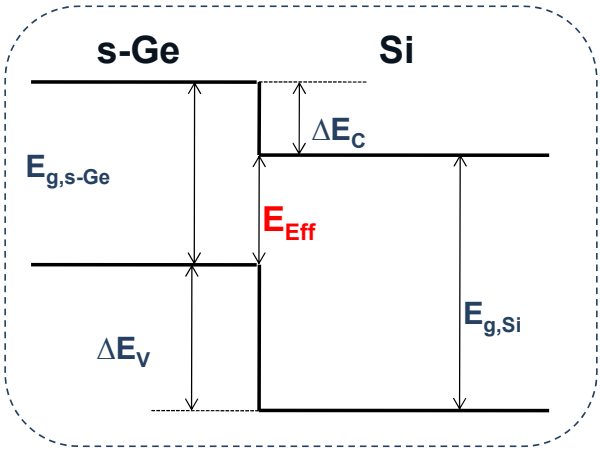


Figure 6.10 Energy band diagram of biaxially strained germanium (s-Ge) grown lattice matched to relaxed silicon

[001]	E_{Eff}	$E_{g,s\text{-}Ge}$	ΔE_C	ΔE_V
Van de Walle et al [6.5]	**0.28**	0.56	0.28	0.84
Rieger et al [6.6]	**0.38**	0.6	0.28	0.74
Yang et al [6.7]	**0.41**	0.71	0.30	0.71
Schwartz et al [6.8]	n/a	n/a	n/a	0.74±0.13
Colombo et al [6.9]	n/a	n/a	n/a	0.74
Band Structure Lab [6.10]	**0.31**	0.41	0.10	0.82

[Units: eV]

Table 6.2 Band alignment of strained germanium on silicon

	[001]	[110]	[111]
$E_{g,s\text{-}Si}$	0.536	0.999	1.128
E_{Eff}	**0.360**	**0.610**	**0.660**

Table 6.3 Orientation dependent band alignment of s-Si on Ge [6.10]

	[001]	[110]	[111]
$E_{g,s\text{-}Ge}$	0.406	0.460	0.650
E_{Eff}	**0.306**	**0.440**	**0.540**

Table 6.4 Orientation dependent band alignment of s-Ge on Si [6.10]

6.4.1 Vertical Pocket Strained Hetero-structure TFET

A vertical pocket strained hetero-structure TFET utilizing the small E_{Eff} is proposed and shown in Figure 6.11. The small E_{Eff} will allow for high tunneling probability and since the tunneling is designed to occur vertically aligned to the gate electric field, a more uniform tunneling with a sharp turn-on can be achieved. The tunneling area can also be increased with increased overlap of pocket to source. The complementary devices require different source (substrate) material for the required hetero-structure.

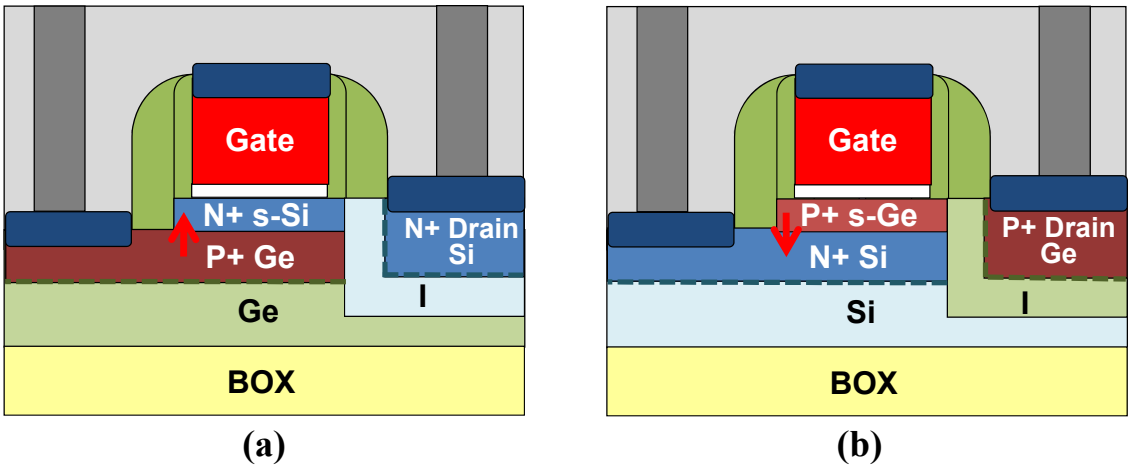


Figure 6.11 Vertical pocket strained hetero-structure **(a)** NTFET **(b)** PTFET

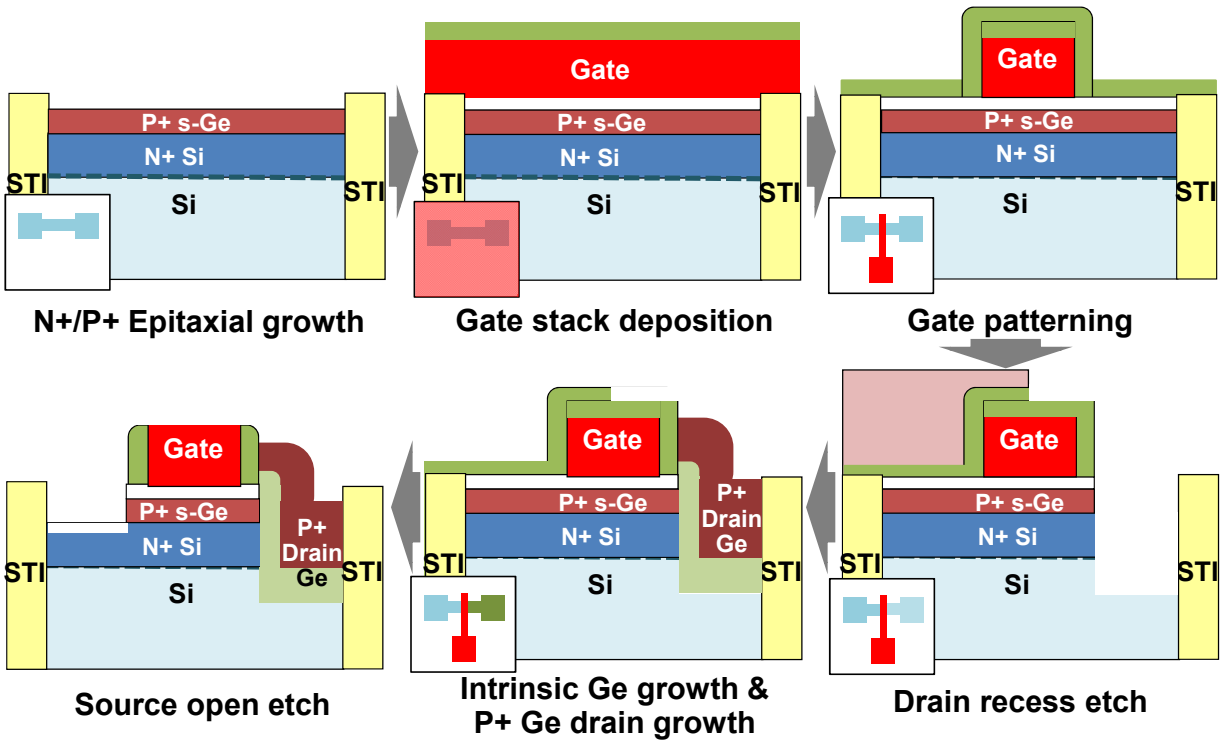


Figure 6.12 Proposed fabrication process flow of vertical pocket hetero-structure PTFET

Figure 6.12 shows the proposed fabrication process flow for the vertical pocket hetero-structure PTFET. PTFET is chosen to allow for processing on bulk STI isolated wafers at Applied Materials. First an N+ source is formed by way of implant and anneal in the defined active regions. Then a thin (2~3nm) of in-situ doped P+ (~$8\times10^{19}cm^{-3}$) germanium pocket is selectively grown. The thin germanium layer needs to be fully strained to achieve the low E_{Eff}. The gate stack is deposited and patterned with the nitride gate hard mask intact. This hard mask is left to protect the gate from being etched during the drain recess etch and regrowth. Then another thin (~10nm) of silicon nitride is deposited to protect the source region. A half mask is used to expose the drain region and etch away the P+ Ge and N+ Si layer. Afterwards a two-step Ge regrowth of undoped followed by P+ in-situ doped is done. The undoped layer is to separate the P+ drain from directly connecting to the pocket. This can cause control of the tunneling between the P+ Ge pocket and N+ source by the drain electric field. Then the source side is opened by etching away the nitride and P+ Ge in the source region is removed to allow for an N+ source contact.

The proposed process flow requires the drain etch to be done aligned to the edge of the gate. But since the gate needs to be protected during this etch, an un-gated pocket region has to exist equal to the thickness of the seal nitride. This added to the undoped germanium epi thickness of the drain will act as a series resistance of the device and when gate fringing field is not effective can form a potential barrier eliminating the sub-60mV/dec subthreshold swing achieved from the band-to-band tunneling between the source and pocket. Figure 6.13 illustrates the effect of the sidewall spacer thickness on the device performance. MEDICI simulation results of all silicon vertical pocket TFETs with varying spacer thickness shows that as the spacer thickness is increased a reduction in the drive current as well as a degradation in the subthreshold swing is to be expected [6.11]. Simulation results suggest that the spacer thickness to be limited to be below 10nm. A seal nitride of 10nm deposited thickness will give us the resulting thickness to be lower than 10nm. Also, this thickness should be sufficient to protect the source during drain recess etch.

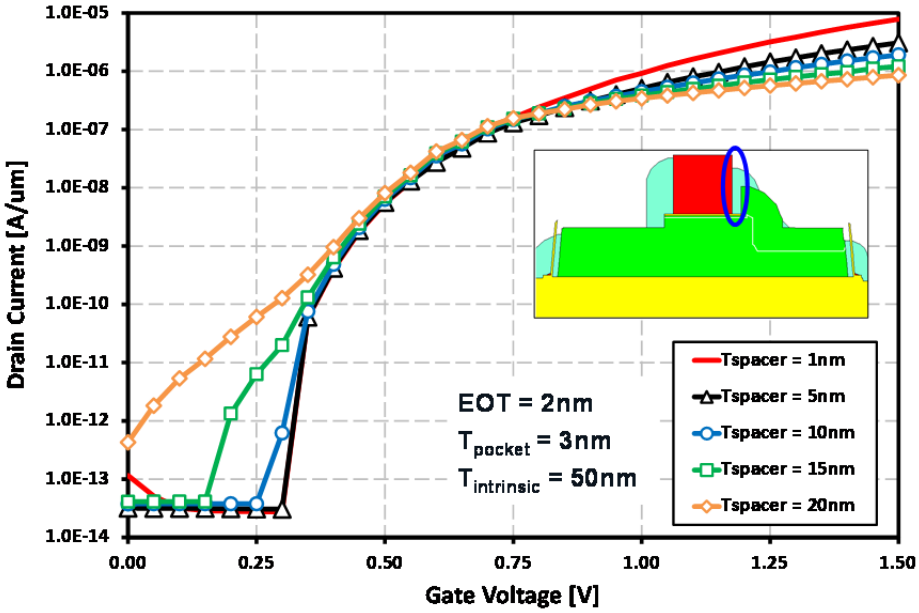


Figure 6.13 Simulated I_D-V_G of vertical pocket TFET with varying spacer thickness

6.4.2 Lateral Strained Hetero-structure TFET

Although the vertical pocket has the advantage of sharp turn-on characteristics due to uniform turn-on of the pocket and increased tunneling area, careful examination of the effective mass in the vertical tunneling direction and energy quantization due to quantum confinement suggests vertical tunneling might not be the optimal tunneling direction in these strained hetero-structure devices.

In the case of strained silicon on germanium, the degeneracy in the conduction band of the silicon split such that the lower bandgap is governed by the 2-fold bands and gives an effective mass in the vertical direction to be the large longitudinal mass ($m_l \approx 0.95m_0$) compared to the lower transverse mass ($m_t \approx 0.19m_0$). This will lower the tunneling probability since it is also exponentially dependent on the effective mass. The vertical tunneling direction is also not favorable due to the energy quantization due to quantum confinement in the vertical direction which will in effect increase E_{Eff} with step function like density of states.

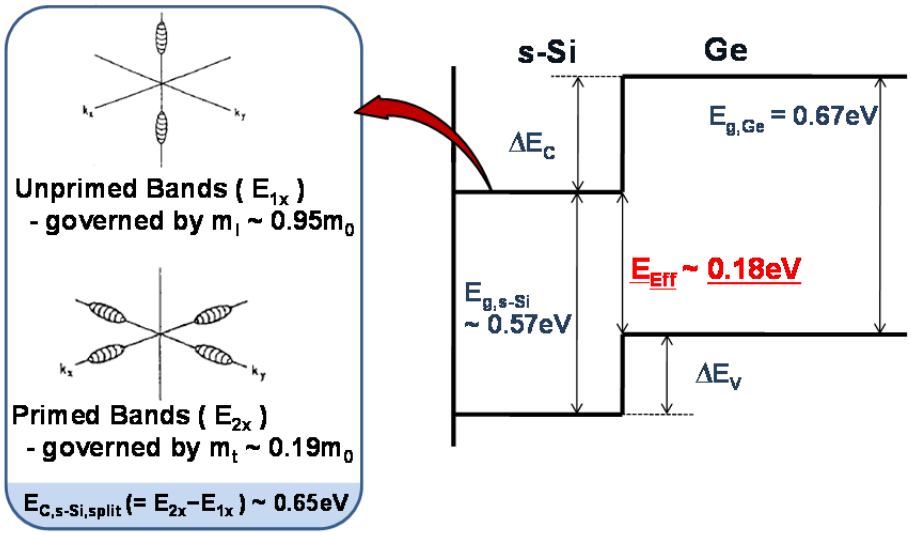


Figure 6.14 Effective mass of biaxially strained silicon on germanium

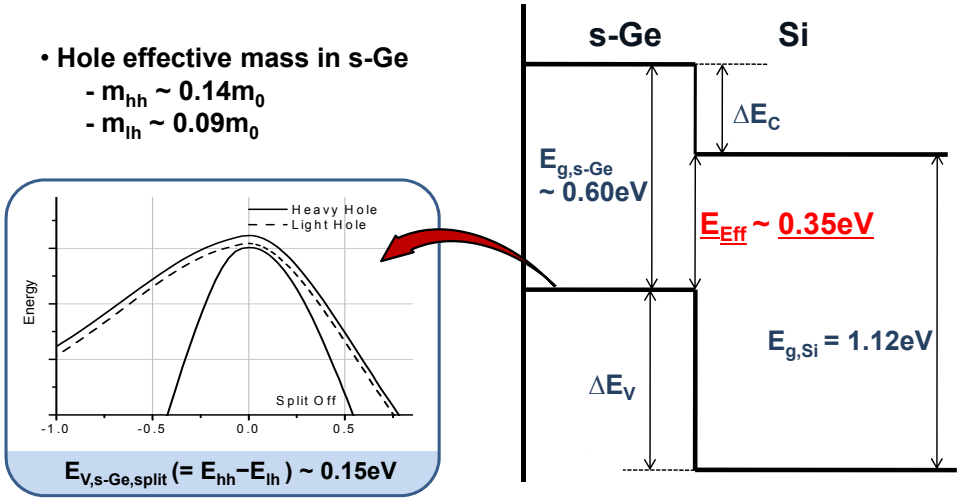


Figure 6.15 Effective mass of biaxially strained germanium on silicon

In the case of strained germanium on silicon, the effective mass problem is not as severe since the heavy and light hole masses do not differ by much ($m_{hh} \approx 0.14m_0$, $m_{lh} \approx 0.09m_0$). But the effect of increase in E_{Eff} still exists in this case as well due to energy quantization from quantum confinement.

With these design considerations in mind, a lateral strained hetero-structure TFET is proposed and is shown in Figure 6.16. These structures allow lateral tunneling to occur between the source material and the strained channel layer. This gives the benefit of the small E_{Eff} as well as the small transverse effective mass for the NTFET and although both structures are still subject to an increase in E_{Eff} due to energy quantization, they do not suffer from the reduction in density of states.

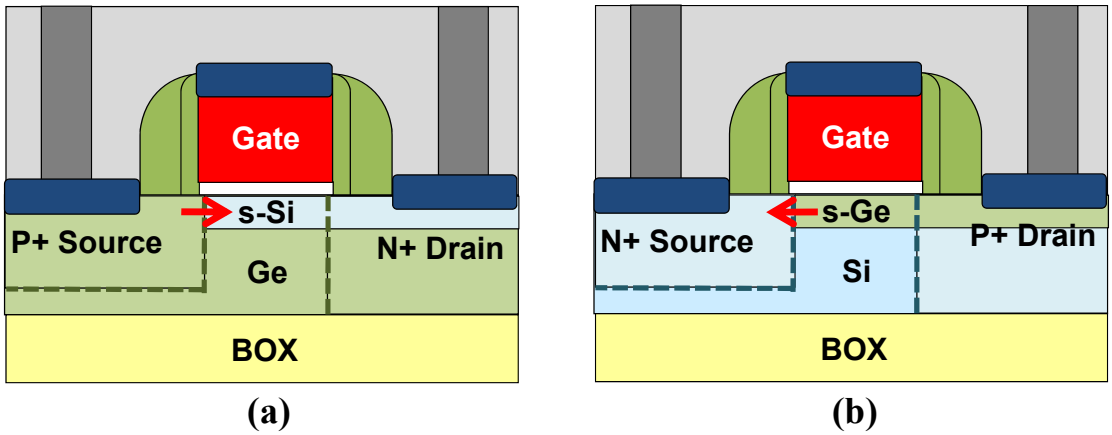


Figure 6.16 Lateral strained hetero-structure **(a)** NTFET **(b)** PTFET

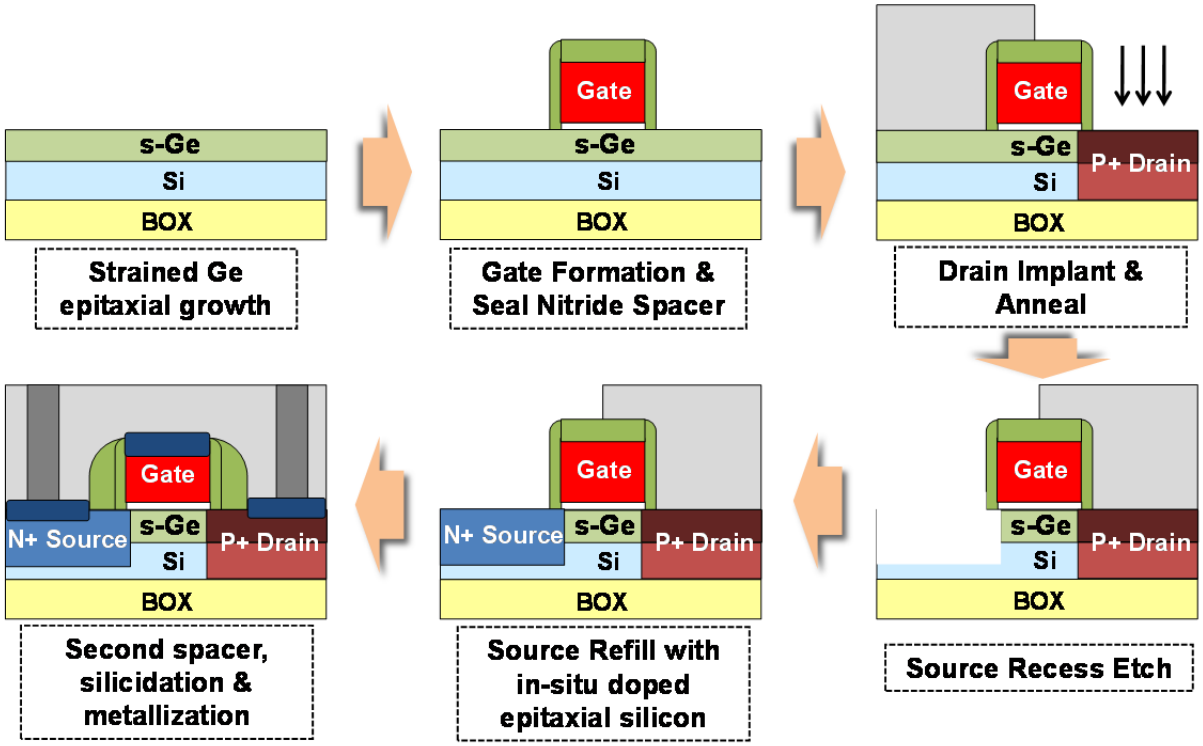


Figure 6.17 Proposed fabrication process flow for later hetero-structure PTFET

The structures are very similar to the non-strained Si/Ge hetero-structure TFETs discussed in previously. As a matter of fact, the P-type device is exactly same as the germanium channel hetero-structure TFET with the only difference being that the germanium layer is very thin and fully strained. Figure 6.17 shows the proposed process flow for the lateral strained hetero-structure PTFET. In-situ doping is suggested to be used during the growth process of the silicon source so that the hetero-junction aligns with the dopant junction.

6.4.3 Strained Germanium Epitaxial Growth Experiments

In order to fabricate the vertical pocket or lateral strained hetero-structure TFET, a fully strained epitaxial layer of germanium need to be grown on silicon substrates with good quality. But due to the 4% lattice mismatch between silicon and germanium, the critical thickness of strained germanium on silicon is less than a few nanometers. Also, the fact that germanium has a tendency to grow in islands is a challenge in growing uniform, good quality, thin germanium films.

Blanket germanium epitaxial growth experiments were carried out at Applied Materials in order to find optimal growth conditions for a fully strained germanium on silicon. Table 6.5 shows the germanium growth experimental splits.

	1	2	3	4	5
Silicon Epi Thickness			**30nm** ($Si_{0.8}Ge_{0.2}$)		
Germanium Epi Thickness	**4nm**	**30nm**	**10nm**	**5nm**	**3nm**
Silicon Cap Epi Thickness	**4nm**		**2nm**	**2nm**	**1nm**

Table 6.5 Germanium blanket epitaxial growth experiment splits

In addition to the challenges in growth, due to the layers being so thin, characterization is also a challenge. XRD (X-Ray Diffraction) measurements were able to be used for thick growth samples such as sample #2 and #3 in the above split table but XRD signal could not be detected on thin growth samples. TEM can be used for a definitive confirmation on the quality of the deposited layers but has the short fall of very long turnaround time and is not ideal to be used when developing and tuning a recipe. So, XRF (X-Ray Fluorescence) measurement recipes were developed for a non-destructive and quick thickness measurement of thin germanium film on silicon. But a controlled growth rate could not be achieved. Further TEM analysis confirmed that thin germanium growth layers (<5nm) causes non-uniform growth with growth in islands and do not form a continuous film until a certain thickness. Figure 6.18 shows the TEM images of germanium growth samples.

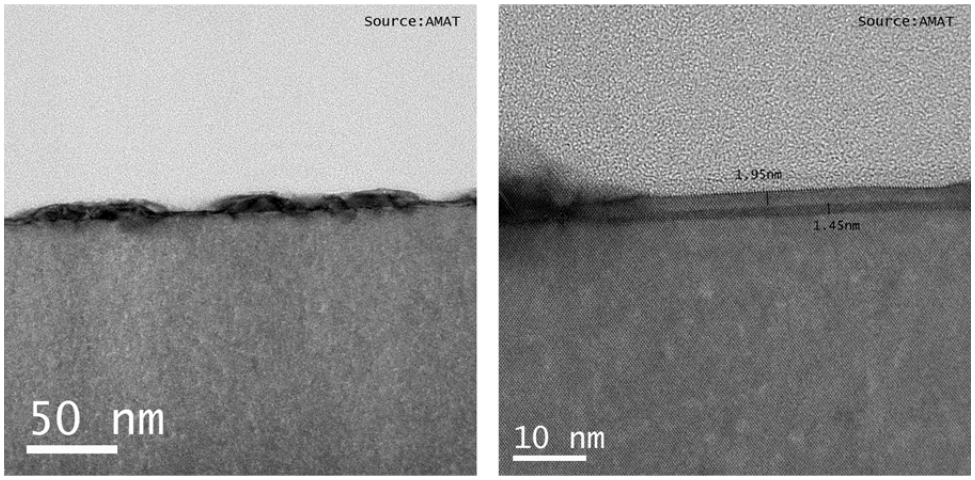


Figure 6.18 TEM images of germanium epitaxial growth sample

6.5 Summary

In this chapter, TFET structures utilizing the lower bandgap of germanium and biaxial strain were examined. Germanium P-I-N TFETs were fabricated on SOI wafers with epitaxially grown relaxed germanium and were shown to have increased tunneling currents compared to silicon P-I-N. But due to the lower bandgap, a much larger increase in the OFF state leakage was observed. The germanium source TFET was explored to lower the leakage current while sustaining the higher tunneling characteristics of germanium. Isotropic etch and regrowth of germanium in the source processes were developed for the fabrication but actual devices shows high gate leakage currents due to gate dielectric damage during the source etch. More development on the source recess etch and gate dielectric damage is required to improve the fabrication process.

Germanium channel TFETs were also fabricated as the complementary device to the germanium source TFET. Devices were fabricated on SOI wafers with relaxed germanium and silicon cap. Selective etch recipes were developed to remove the germanium at the source region with undercut. Poor selectivity during the regrowth of silicon in the source caused an electrical short between the gate and source thus showing high gate leakage currents. More development is required to improve the selectivity of silicon growth on silicon to silicon nitride.

Biaxially strained Si/Ge hetero-structures are studied and shown to be able to achieve very low effective tunneling bandgap of ~0.18eV for s-Si on Ge and ~0.3eV for s-Ge on Si. Orientation dependent simulation of the biaxial strain shows lowest E_{Eff} can be achieved on [001] surface. Vertical pocket and lateral TFET structures utilizing the low E_{Eff} of strained Si/Ge are proposed. Simulation of the vertical pocket TFET shows seal nitride spacer thickness is crucial in the performance for the suggested process flow. Lateral strained hetero-structure TFET designs benefit from a lower electron effective mass in the tunneling direction.

In order to fabricate the strained hetero-structure TFETs, epitaxial growth of strained germanium experiments were conducted. Due to its nature of growing on islands, thin epitaxial growth of strained germanium on silicon is challenge.

6.6 References

[6.1] H. Shang, M. M. Frank, E. P. Gusev, J. O. Chu, S. W. Bedell, K. W. Guarini, M. Ieong, "Germanium channel MOSFETs: Opportunities and challenges," *IBM Journal of Research and Development*, vol.50, no.4.5, pp.377-386, July 2006

[6.2] C.O. Chui, H. Kim, D. Chi, P.C. McIntyre, and K.C. Saraswat, "Nanoscale germanium MOS Dielectrics-part II: high K gate dielectrics," *IEEE Transactions on Electron Devices*, vol.53, no.7, pp.1509-1516, July 2006

[6.3] A. Satta, E. Simoen, R. Duffy, T. Janssens, T. Clarysse, A. Benedetti, M. Meuris, and W. Vandervorst, "Diffusion, activation, and regrowth behavior of high dose P implants in Ge", *Applied Physics Letters*, vol. 88, pp. 162118, 2006

[6.4] R. Li, H. B. Yao, S.J. Lee, D. Z. Chi, M.B. Yu, G.Q. Lo and D. L. Kwong, "Metal germanide Schottky Source/Drain transistor on Germanium substrate for future CMOS technology", *Thin Solid Films*, vol. 504, Issues 1-2, 10 May 2006, pp. 28-31

[6.5] C. G. Van der Walle, R. M. Martin, "Theoretical calculations of heterojunction discontinuities in the Si/Ge system," *Physical Review B*, vol. 34, no. 8, pp. 5621-5634, October 1986

[6.6] M. M. Rieger and P. Vogl, "Electronic-band parameters in strained $Si_{1-x}Ge_x$ alloys on $Si_{1-y}Ge_y$ substrates," *Physical Review B*, vol. 48, no. 19, pp. 14276-14287, November 1993

[6.7] L. Yang, J. R. Watling, R. C. W. Wilkins, M. Boriçi, J. R. Barker, A. Asenov and S. Royang, "Si/SiGe heterostructure parameters for device simulations," *Semiconductor Science and Technology*, vol. 19, no. 10, pp. 1174-1182, August 2004

[6.8] G. P. Schwartz, M. S. Hybertsen, J. Bevk, R. G. Nuzzo, J. P. Mannaerts, and G. J. Gualtieri, "Core-level photoemission measurements of valence-band offsets in highly strained heterojunctions: Si-Ge system," *Physical Review B*, vol. 39, no. 2, pp. 1235-1241, January 1989

[6.9] L. Colombo, R. Resta, and S. Baroni, "Valence-band offsets at strained Si/Ge interfaces," *Physical Review B*, vol. 44, no. 11, pp. 5572-5579, September 1993

[6.10] A. Paul, M. Luisier, N, Neophytou, R. Kim, J. Geng, M. McLennan, M. Lundstrom, and G. Klimeck, "Band Structure Lab," http://nanohub.org/resources/bandstrlab, 2011

[6.11] Taurus MEDICI User Manual, Version A-2007.12, Synopsys, Mountain View, CA

Chapter 7

Conclusions

7.1 Summary of Work

This work has focused on researching the design and fabrication of band-to-band tunneling transistors as an alternative to MOSFETs for low power applications.

In Chapter 2, the band-to-band tunneling phenomenon was modeled considering the transition of the effective masses in the two bands. This model was then applied to 2D device simulator MEDICI to simulate germanium TFETs. Default tunneling model parameters for silicon are shown to match well with calculated results.

In Chapter 3, a baseline fabrication process flow for the silicon P-I-N structure TFET on SOI wafers was developed with LOCOS isolation on SOI and protruding high-k foot processes to reduce gate leakage. TFET electrical characterization schemes are discussed for current and capacitance measurements. Implant energy splits show degradation of subthreshold swing due to larger straggle and SOI thickness splits show better gate control with thinner SOI results in improved subthreshold characteristics. Similar gate length dependence to drive current of MOSFETs were observed in P-I-N TFETs.

In Chapter 4, the silicide source TFET is discussed. Dopant segregation with nickel silicide is used to form abrupt source profile and wedge shaped pocket which is shown to be critical to the performance of the device through simulation. A subthreshold swing of 46mV/dec was measured at the steepest region with an I_{ON}/I_{OFF} ratio of 7×10^7 for 1V operation which is the highest reported for silicon TFETs. Temperature dependence measurements of the subthreshold swing was shown to have nonlinear relation. Screening criteria to eliminate falsely low subthreshold swing resulting from gate leakage, zero crossing or transient effects is developed and applied. It was shown that without the screening applied a much larger portion of devices would seem to have sub-60mV/dec subthreshold swing. Selective nickel silicide using germanium shows slower silicidation of SiGe and a structure using this process is proposed.

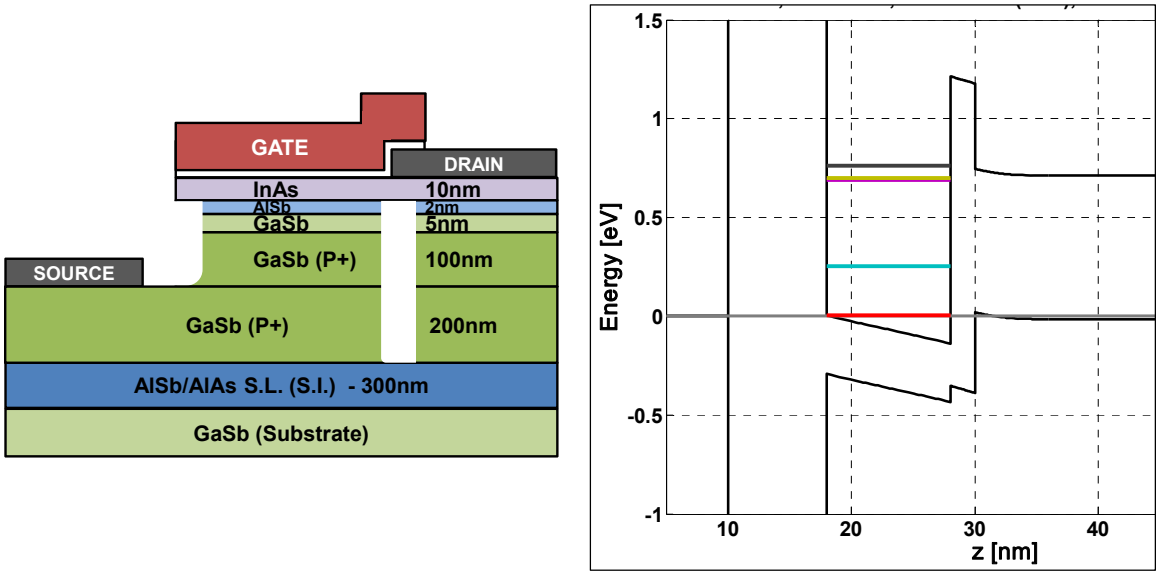
In Chapter 5, P-I-N TFETs with intentional gate to source overlap and undoped vertical pocket TFETs were studied. High tunneling currents of >100μA/μm at an overdrive of 2.5V were observed in intentional overlapped TFETs. The spike+flash combination was shown to improve the performance through higher dopant activation. NDR behavior was observed in the

diode measurements confirming the current mechanism as tunneling. Undoped vertical pocket TFETs with various pocket thicknesses were fabricated but due to a failure in the silicon epitaxial growth process, improvement could not be observed between control and pocket TFETs.

In Chapter 6, TFET structures utilizing the lower bandgap of germanium and biaxial strain were examined. Germanium P-I-N TFETs show improved drive current compared to the silicon counterpart but a higher increase in OFF state current caused worse subthreshold swing. Germanium source TFETs were explored to reduce this OFF state current but gate dielectric damage during source recess etch process caused high gate leakage and band-to-band tunneling current extraction impossible. Germanium channel TFETs were also fabricated as the complementary device to the germanium source TFET. Poor selectivity during source regrowth bridges the source to gate and thus high gate currents. Biaxially strained Si/Ge hetero-structures can achieve a very low E_{Eff} of ~0.18eV for s-Si on Ge and ~0.3eV for s-Ge on Si. Vertical pocket and lateral TFET structures utilizing this low E_{Eff} of strained Si/Ge were proposed. Lateral strained hetero-structure TFET designs benefit from a lower electron effective mass in the tunneling direction. Epitaxial growth experiments show thin growth of germanium on silicon is a challenge.

7.2 Future Directions

In a means to increase the drive current of TFETs, III-V materials are being studied extensively [7.1~7.5]. III-V materials with its wide selection of material combinations and well developed MBE growth techniques allow for a new knob in the design of the TFET. InAs/AlSb/GaSb NDR diodes which have been reported to possess high PVCR [7.6] can be utilized in the following quantum well TFET structure.



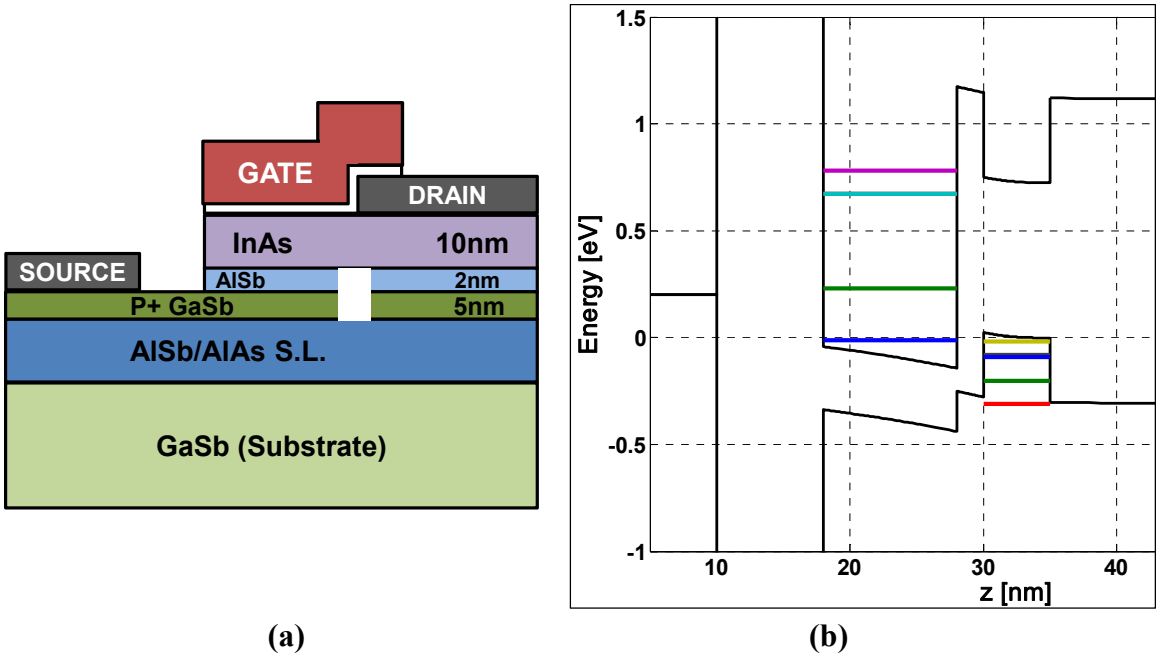
(a) **(b)**

Figure 7.1 III-V quantum well TFET **(a)** Structure **(b)** Energy band diagram

Figure 7.1 shows the III-V quantum well TFET structure and simulated energy band diagram using Nextnano [7.7]. The broken gap between the InAs and GaSb allows for a high tunneling probability and can be controlled with the thickness of AlSb and the position of the ground state energy of the quantum well formed by the gate dielectric and AlSb. Tunneling into the quantized energy state will show sharp switching behavior. The initial increase in the tunneling current can be tuned by controlling the ground state energy level with the quantum well thickness (T_{InAs}). The higher the ground state energy is the larger the density of states and thus a larger initial jump in tunneling current. High leakage currents observed in III-V TFETs [7.1] can be reduced by isolating the drain contact region from the source and by using a AlSb/AlAs super-lattice for the semi-insulating layer.

Double quantum well designs where the tunneling occurs from one quantum well to another can be also designed using the same material combination of InAs/AlSb/GaSb. Figure 7.2 shows the double quantum well TFET structure and energy band diagram. The GaSb layer thickness is much thinner (5nm) in this case to form a quantum well in the valence band between the AlSb and AlSb/AlAs super-lattice. Landing a source contact on the thin GaSb layer will be a challenge in fabricating this structure. A sharper turn-on characteristics is expected from the tunneling occurring between the ground state energy of valence band quantum well in GaSb to the ground state energy of the conduction band quantum well in InAs. Also, the initial jump in tunneling current is expected to be higher than the single well case due to the overlap of energies occurring at a higher initial density of states due to quantization.

III-V TFETs using InAs/AlSb/GaSb material combination shows to have a high potential in improving the performance of band-to-band tunnel transistors.



(a) **(b)**

Figure 7.2 Double quantum well TFET **(a)** Structure **(b)** Energy band diagram

7.3 References

[7.1] S. Mookerjea, D. Mohata, R. Krishnan, J. Singh, A. Vallett, A. Ali, T. Mayer, V. Narayanan, D. Schlom, A. Liu, and Datta, "Experimental demonstration of 100nm channel length $In_{0.53}Ga_{0.47}As$-based vertical inter-band tunnel field effect transistors (TFETs) for ultra low-power logic and SRAM applications," *IEDM Technical Digest,* pp. 949-951, Dec. 2009

[7.3] J. Knoch and J. Appenzeller, "Modeling of High-Performance p-Type III–V Heterojunction Tunnel FETs," *IEEE Electron Device Letters*, vol. 31, no. 4, pp. 305-307, April 2010

[7.3] W. Lingquan, E. Yu, Y. Taur, and P. Asbeck, "Design of Tunneling Field-Effect Transistors Based on Staggered Heterojunctions for Ultralow-Power Applications," *IEEE Electron Device Letters*, vol. 31, no. 5, pp. 431-433, May 2010

[7.4] G. Dewey, B. Chu-Kung, J. Boardman, J.M. Fastenau, J. Kavalieros, R. Kotlyar, W.K. Liu, D. Lubyshev, M. Metz, N. Mukherjee, P. Oakey, R. Pillarisetty, M. Radosavljevic, H.W. Then, R. Chau, "Fabrication, characterization, and physics of III–V heterojunction tunneling Field Effect Transistors (H-TFET) for steep sub-threshold swing," *IEDM Technical Digest*, pp. 33.6.1-33.6.4, Dec. 2011

[7.5] R. Li, Y. Lu, S. D. Chae, G. Zhou, Q. Liu, C. Chen, M. Shahriar Rahman, T. Vasen, Q. Zhang, P. Fay, T. Kosel, M. Wistey, H. Xing, S. Koswatta, and A. Seabaugh, "InAs/AlGaSb heterojunction tunnel field-effect transistor with tunnelling in-line with the gate field," *Physica Status Solidi C*, 9, no. 2, pp. 389–392, 2012

[7.6] J. N. Schulman and D. H. Chow, "Sb-Heterostructure Interband Backward Diodes," *IEEE Electron Device Letters*, vol. 21, no. 7, pp. 353-355, July 2000

[7.7] Nextnano software available online at http://www.nextnano.de